

## DOCUMENT RESUME

ED 378 076

SO 024 152

AUTHOR Finlayson, Sandra June  
TITLE A Critical Description of Selected State Level Visual Arts Assessment Instruments.  
PUB DATE Jun 88  
NOTE 103p.; Master's Thesis, University of Oregon.  
PUB TYPE Dissertations/Theses - Masters Theses (042)  
  
EDRS PRICE MF01/PC05 Plus Postage.  
DESCRIPTORS \*Art Education; \*Educational Assessment; Elementary Secondary Education; \*Evaluation Methods; \*Evaluation Research; Holistic Evaluation; Measurement Techniques; Standardized Tests; \*State Standards; \*Visual Arts  
IDENTIFIERS State Assessment of Educational Progress

## ABSTRACT

Statewide assessments, developed by several states, of student learning in the visual arts, copies of their visual arts assessment tests, technical manuals, supporting documentation, and evaluative approaches other than testing were critically described within this study. Information gathered resulted in grouping the descriptions of state level visual arts assessments into three categories: (1) states that indicated visual arts assessments did not exist or existed in an abbreviated fashion; (2) states that indicated substantial research and development of visual arts questioning, tests, or alternative approaches to student evaluation; and (3) states that sent visual arts assessment tests. Analysis of the materials received put forth the following conclusions: (1) The forms of state visual art tests ranged from non-existent to holistic to traditional assessment testing instruments with a slant toward representation of art education values. (2) The visual arts assessment instruments in the form of tests provided a sampling of student learning but were neither comprehensive nor disclosing of many aspects of art classroom learning. The materials measured certain specific cognitive outcomes, affective responses and attitudes, and production characteristics in a delineated and timed setting. (3) A thoughtful evolutionary process of viable assessment instruments was planned that considered classroom teacher ideas and the existent state level curriculum. (4) The testing instruments utilized techniques and practices valued by the test and measurement fields. (5) The holistic evaluative approaches to student learning experiences served as the most extensive sample of alternative techniques. Study results provided a critical description of information received, recommendations for further research regarding evaluation of student learning in the visual arts, and implications for art education which included the recognition of three issues: (1) the role of evaluation in the evaluation setting; (2) the need for dialogue among art educators; and (3) a clear recognition that differences exist between specified and unspecified outcomes. The document concludes with the vitae of the author, four appendixes that include an outline of the test evaluation, and a bibliography.  
(MM)

A CRITICAL DESCRIPTION OF SELECTED STATE  
LEVEL VISUAL ARTS ASSESSMENT  
INSTRUMENTS

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

☒ This document has been reproduced as  
received from the person or organization  
originating it  
☐ Minor changes have been made to improve  
reproduction quality

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official  
OERI position or policy

"PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

SANDRA  
FINLAYSON

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC) "

by

SANDRA JUNE FINLAYSON

A THESIS

Presented to the Department of Art Education  
and the Graduate School of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Master of Arts

June 1988

SO 024 152

APPROVED: \_\_\_\_\_  
Dr. Jane Maitland-Gholson

## An Abstract of the Thesis of

Sandra June Finlayson                      for the degree of                      Master of Arts  
in the Department of Art Education                      to be taken                      June 1988

Title: A CRITICAL DESCRIPTION OF SELECTED STATE LEVEL VISUAL  
ARTS ASSESSMENT INSTRUMENTS

Approved: \_\_\_\_\_  
Dr. Jane Maitland-Gholson

Statewide assessment of student learning in the visual arts had been reportedly developed by several states. Among them were Connecticut, Delaware, Hawaii, Louisiana, Maryland, Michigan, Minnesota, Missouri, Pennsylvania, South Carolina, Utah, Wisconsin, and the District of Columbia. Copies of visual arts assessment tests, technical manuals, supporting documentation, and evidence of evaluative approaches other than testing were gathered to be critically described within this study.

Information gathered resulted in the grouping of the descriptions of state level visual arts assessment into three categories: (1) states whose replies indicated that visual arts assessment did not exist or existed in an abbreviated fashion; (2) states who indicated substantial research and development of visual arts questioning, tests, or alternative approaches to student evaluation; and (3) states which sent visual arts assessment tests.

Study results include a critical description of information received,

recommendations for further research regarding evaluation of student learning in the visual arts, and implications for art education and general education.

© 1988 Sandra June Finlayson

## VITA

NAME OF AUTHOR: Sandra June Finlayson

PLACE OF BIRTH: Lewiston, Idaho

DATE OF BIRTH: June 1, 1950

## GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon  
Idaho State University  
University LaVal, Quebec  
University of Idaho

## DEGREES AWARDED:

Master of Arts, 1988, University of Oregon  
Bachelor of Arts, 1972, University of Idaho

## AREAS OF SPECIAL INTEREST:

Assessment of Student Learning in the Visual Arts and General  
Education

Batik Manufacturing with Technological Advancements

Teaching Methodology Supporting Individual Inquiry

## PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, Department of Art Education, University of  
Oregon, Eugene, 1986-1988

Art Teacher, School District 55, Blackfoot, Idaho, 1980-1986

Teacher and Coordinator, Bingham County Community Education,  
Blackfoot, Idaho, 1977-1980

General Educator, Shoshone-Bannock Tribal School, Fort Hall, Idaho,  
1975-1976

**PRESENTATIONS:**

"Can We Speak the Language? Toward an Understanding of  
Standardized Testing Techniques and Alternatives", National Art  
Education Association Conference, Los Angeles, California, April,  
1988



## TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION . . . . .	1
	Statement of the Problem . . . . .	1
	Need for the Study . . . . .	2
	Purpose of the Study . . . . .	4
	Research Questions . . . . .	5
	Definition of Terms . . . . .	5
II.	REVIEW OF LITERATURE . . . . .	9
	Evaluation of Learner Achievement in the Visual Arts . . . . .	9
	Standardized Testing of Learner Achievement in the Visual Arts . . . . .	21
	Criteria Applied to Standardized Tests . . . . .	25
III.	DESIGN OF THE STUDY. . . . .	27
	Data to be Gathered . . . . .	27
	Methods of Gathering Data . . . . .	28
	Analysis . . . . .	29
IV.	RESULTS . . . . .	32
	States Whose Replies Indicated Visual Arts Assessment Did Not Exist or Existed in an Abbreviated Fashion . . . . .	34
	States Whose Replies Indicate Substantial Research and Development of Visual Arts Questioning, Tests, or Alternative Approaches . . . . .	39
	States Which Sent Standardized Visual Arts Assessment Tests. . . . .	47

V.	DISCUSSION . . . . .	65
	Recommendations for Further Research. . . . .	67
	Implications for Art Education . . . . .	69
	Implications for General Education . . . . .	70
	Closure . . . . .	71
APPENDIX		
A.	LETTER OF REQUEST FOR DATA . . . . .	73
B.	LETTER OF SUPPORT FROM RICHARD LATOUR. . .	76
C.	SECOND LETTER OF REQUEST FOR DATA. . . . .	78
D.	ANASTASI'S "SUGGESTED OUTLINE FOR TEST EVALUATION" . . . . .	80
	BIBLIOGRAPHY . . . . .	84

## LIST OF TABLES

Table	Page
1. Record of Responses to Requests for State Level Visual Arts Assessment Instruments and Supporting Documentation . . . . .	33

## CHAPTER I

### INTRODUCTION

Traditionally, within the field of art education, tests have focused on artistic aptitude and ability (Educational Testing Service, 1987). A more recent development has been standardized visual arts assessment tests which are administered to various grade levels by certain states within their public school systems ("States Move", 1986).

#### Statement of the Problem

It has been reported that twelve states and the District of Columbia have developed or are in the process of developing testing instruments which would measure student achievement in the visual arts (Council of Chief State School Officers, 1985; "States Move", 1986). In the reporting of the existence of state level assessment testing in the visual arts, no mention has been made of the instruments having been reviewed using criteria from the tests and measurements field for evaluating standardized tests. A comprehensive study which described these assessment tests and their development or status of use was not found prior to the initiation of this study.

Reporting on art assessment testing (National Art Education Association, 1986) has been diverse in content. Articles, while informative, have been written by individual representatives of state departments of education, leaders in art education, fine arts specialists, or assessment specialists. The choice and availability of information about the tests has varied from state to state. It was found that the Educational Testing Service had no copies of state level visual arts assessment tests in their test collection nor did their service have any reviews on record of the aforementioned tests.

In order to provide a clearer picture of the development and use of state level visual arts assessment tests, this researcher felt it important to develop a body of information which included descriptions of actual tests using criteria accepted by experts in the test and measurement field.

### Need for the Study

Commentary from leaders in the arts and art education underscores the need for awareness of the role of the standardized test as an evaluative instrument in visual arts education. Hodsoll, chair of the National Endowment for the Arts, in reference to state level standardization of testing in the arts, has made the forceful comment: "Ways must be found for such evaluation to proceed on a continuing basis; the fact that we have less experience with 'testing' in the arts than in other academic areas is no excuse for not finding appropriate ways to do it." (1985, p. 249).

Art educators have continued to mold and articulate the potential

structuring of the Discipline-Based Art Education movement (DBAE). The Discipline-Based Art Education movement, as the basic content for curriculum development, promotes the overarching inclusion of four parent disciplines within the visual art education setting: Aesthetics, Art Criticism, Art History, and Art Production. Exploration of aesthetic valuing, development of visual critical skills, an awareness of art historical perspectives, and art production through skill development have been put forth as invaluable curriculum components.

Educational leaders have reinforced the importance of the role of evaluation as a process for improving the quality of curriculum, teaching , and learning (Eisner, 1987). Information has appeared in the literature which has supported an increased awareness for art educators of evaluation and measurement of student achievement in the visual arts (Day, 1985; Hoepfner, 1984; Pratt, 1983; "States Move", 1986).

The pioneering efforts of the National Assessment of Educational Progress in visual arts assessment in the first and second national assessment of art (conducted in 1974-75 and 1978-79) provided a model for the development of art objectives with suitable testing exercises, utilization of assessment procedures, and test modification (Pratt, 1983). Even with such thorough and well-grounded research which developed techniques and questioning strategies considered suitable for the measurement of certain learner progress in the visual arts (Wilson, 1971), art educators have reportedly shown resistance to standardization (Day, 1985; Hamblen, 1987). Crossing over into the test and measurements arena has, for many, represented

uncharted territory. However, aversion to standardization alone has not provided art educators with the tools to sensitively protect and advocate valued positions. Rather, information exchange, awareness of techniques, and learning from the work of others are needed in order to articulate a concern.

It was to this end of increased understanding of visual arts assessment instruments and shared information among art educators, state level arts administrators, general educators, and assessment specialists, that this study was undertaken.

#### Purpose of the Study

The purpose of this study was to identify, gather, and describe state level visual arts assessment tests of learner progress in order to provide a body of information to which art educators, state level arts administrators, general educators, and assessment specialists may refer to determine the characteristics of certain state level visual arts assessment instruments which have been developed and used. Availability of such information is intended to assist art educators with possible comparisons, development, or reworking of assessment tools for future measurement of learner progress in the visual arts. The study also included a secondary, exploratory purpose which involved identifying alternatives to standardized testing in the visual arts recommended by some states.

### Research Questions

This researcher has addressed the following questions by identifying, gathering, and describing information about reported state level visual arts assessment tests.

1. What form have the state visual arts tests taken?
2. What do these instruments measure?
3. Does it appear that state level curriculum guides or learning objectives in the visual arts were taken into consideration in test construction?
4. Did the tests' characteristics correspond to characteristics considered to be important by experts in the test and measurement field?
5. What alternative techniques for evaluating student progress in the visual arts surfaced at the state level?

### Definition of Terms

The following terms have been used in various portions of the study. The definitions were taken from recent art education literature which focused on evaluation of student learning (Day, 1985; Hoepfner, 1984), a text on psychological testing (Anastasi, 1982), and a chapter on evaluating student learning within an art education text (Eisner, 1972).

Norm: A norm represents a standard of average performance of people of similar age and background. In the process of standardizing a test, it is administered to a large, representative sample of the type of persons for whom it is designed. This group, known as the standardization sample, serves to establish the norms (Anastasi, 1982, p. 25).



Sample: The sample refers to the group of persons actually tested in order to establish norms. An effort is made to obtain a representative cross section of the population with reference to geographical distribution, socioeconomic level, ethnic composition, and other relevant characteristics. Relevance of the norms is closely linked to representativeness of the sample, and initial clarity in the definition of the specific population to which the norms are to apply helps to prevent sampling error (Anastasi, 1982, pp. 86-87).

Reliability: A measure of reliability characterizes the test when administered under normal conditions and given to persons similar to those constituting the normative (standardization) sample. Reliability indicates the extent to which individual differences in test scores are attributable to 'true' differences in the characteristics under consideration and the extent to which score differences are attributable to chance errors. For example, if similar individuals had like scores on a given test, or if an individual had similar scores on the same test after having been retested, the test would be considered to be reliable (Anastasi, 1982, p. 102).

Validity: The validity of a test concerns what the test measures and how well it does so. It tells us what can be inferred from test scores. Test validity may not be reported as 'high' or 'low' without established reference to the particular use for which the test is being considered (Anastasi, 1982, p. 131).

Content Validity: Content validity represents one category of methods employed for investigating the relationships between performance on a test and other independently observable facts about behavior characteristics under consideration. Content validation involves the systematic examination of test content to determine whether it covers a sample of the behavior domain to be measured. Ideally, the behavior domain is clearly stated and described in initial test specifications. This enables the test constructor to address major aspects in correct proportions, and content validity is built from the base of test specifications (Anastasi, 1982, pp. 131-132).

Face Validity: Face validity refers to what the test appears superficially to measure. This is not to be confused with validations which are determined by technical processes such as content validity. Rather, face validity pertains to whether the test 'looks valid' to those who take the test, decide on its use, and other technically untrained observers. Face validity is a desirable trait, particularly in establishing rapport with examinee (Anastasi, 1982, p.136).

Test specifications: Test specifications are drawn up before any items are prepared. Included should be outlines of both the instructional objectives (types of learning) and the content areas to be covered within the instructional time span to be tested. This information guides development of appropriate items in proper proportions and promotes content validity (Anastasi, 1982, pp. 410-413).

Item analysis: After test construction and piloting, item analysis is used to determine whether an item is actually effective in assessing learning objectives. A desirable item will be able to discriminate among learners. It will not be too easy, too hard, irrelevant, or non-informative. Eliminating certain items will contribute to overall content validity (Hoepfner, 1984, p. 252).

Technical manual: A test's technical manual should provide essential information on administering, scoring, and evaluating the test. It should contain detailed instructions, scoring key, norms, and data on reliability and validity. Ideally, the technical manual enables test users to evaluate the test prior to selecting it for use (Anastasi, 1982, p. 20).

Evaluation: Evaluation is viewed as having the primary mission of securing information which will allow the teacher or curriculum planner to improve the educational process. An educational tool, evaluation serves a diagnostic role in an ongoing process of upgrading the learning process and product. Evaluation may function most effectively when reward or punishment with their accompanying connotations has been dismissed. (Eisner, 1972).

Testing: Testing is one of several vehicles for gathering information which samples student abilities (Eisner, 1972).

Grading: To grade is to assign a symbol which represents a learner's relative achievement with regard to a given set of criteria (Eisner, 1972).

Diagnostic evaluation: Diagnostic evaluation is designed to analyze an individual's specific strengths and weaknesses in a subject and to suggest causes of his or her difficulties (Anastasi, 1982, p. 414).

Formative evaluation: Formative evaluation is evaluation that is used to inform the teacher for the purpose of in-process educational decision-making. Flexibility on the teacher's part is desirable in responding to the need to revise instructional approaches or reteach areas of curriculum which were not understood (Day, 1984, p. 235).

Summative evaluation: Summative evaluation produces information about student learning which allows the teacher to reach summative statements about each student's accomplishments and relative success of the class (Day, 1984, p. 235).

Norm-referenced tests: Two goals of norm-referenced testing are the wide dispersion of outcome scores and a high level of discrimination among examinees with regard to abilities or achievements being measured. Norms will have been established by piloting the test on a representative sample. Norm-referenced tests are linked to the 'normal' distribution of scores,

theoretically without concern for the content of the items (Hoepfner, 1984, p. 252).

Criterion-referenced testing: Criterion-referenced tests of achievement are linked directly to the learning objectives established for the curriculum. No a priori attention is paid to the distribution of resulting scores. Successful completion of criterion-referenced tests is one indicator of mastery of content (Hoepfner, 1984, p. 252).

Many standardized tests incorporate features which are both norm-referenced and criterion-referenced in nature.

Unstructured tests: Unstructured tests in art education may be production tests with no restrictions. A state of maximum unreliability exists in both the examinee's response and the procedures followed in scoring (Hoepfner, 1984, pp. 254-257).

Verbally structured tests: Verbally structured tests in art production provide both the examinee and the scorer with more concise criteria thereby increasing reliability. Verbally structured test items in art criticism and art history generally are exemplified by questioning which leads the examinee to select an image after having been presented with information in words (Hoepfner, 1984, pp. 254-257).

Object structured test: Object structured tests in art production may involve drawing from a model (object) in which all examinees view the same model. Object structured test items in art history and art criticism employ questioning which leads the examinee to select words after having been presented with information in a visual (object) (Hoepfner, 1984, pp. 254-257).

Visual arts assessment tests/instruments: In this document, visual arts assessment tests/instruments will refer to tests or other instruments which have been developed by certain states for the purposes of measuring student achievement in the visual arts in the public school systems, K-12.

## CHAPTER II

### REVIEW OF LITERATURE

In order to critically describe standardized tests and techniques designed to measure student achievement in the visual arts, it became apparent that greater awareness in three topic areas would be of importance to this study: a) the multiplicity of evaluative approaches to learner achievement in the visual arts and current issues regarding their application, b) standardized testing of learner achievement in the visual arts, and c) criteria applied to standardized tests for the purpose of their description from the perspective of the test and measurement field.

#### Evaluation of Learner Achievement in the Visual Arts

In gathering information regarding evaluation of learner achievement in the visual arts, approaches to evaluation of learning which form points of reference for many art educators in the field today were reviewed . Current change within the curricular structure of the art education field and its relatedness to evaluation of student achievement was seen as an important part of the context of the study . Proposed curricular restructuring with more tightly

defined objectives and measurable learner outcomes has brought commentary from art educators, which has identified and shaped the issues. Research on the effect of evaluation on learner motivation and the field study of multiple approaches to evaluation of learner achievement in the arts contributed to the context of the review of literature.

Therefore, this section of the review of literature will include three topics:

a) approaches to evaluation of student learning in the visual arts, b) proposed curricular changes in the visual arts and surrounding issues related to evaluation of learner achievement, and c) evidence of research in specific evaluative techniques in the visual arts.

#### Approaches to Evaluation of Student Learning in the Visual Arts

It was determined that texts representative of handbooks for art educators which included chapters on evaluation (Day, 1985) would be reviewed.

Creative and Mental Growth (Lowenfeld, 1957; Lowenfeld & Brittain, 1987) was included as representative of the position opposing grading of student artwork.

Approaches to Art in Education (Chapman, 1978), Children and Their Art (Gaitskell, Hurwitz, & Day, 1982), and Educating Artistic Vision (Eisner, 1972) provided examples of approaches supporting multiple means of evaluating student learning. An investigation of Brent Wilson's seminal work in synthesizing a variety of questioning techniques appropriate for identifying learning relative to the visual arts (Wilson, 1971) gave additional evidence to the aspect of multiplicity in evaluative approaches.

The Lowenfeld (1956) and Lowenfeld/Brittain (1987) perspective is that of protection of the child/student's burgeoning expressive nature through the stance that "grading in art has no function" (1987, p. 175) and that the imposition of the teacher's values in an evaluative situation are meaningless to the child. In the most recent edition of Creative and Mental Growth (1987) Brittain has cited several studies which support the idea that impending evaluation may negatively effect creative production, levels of student interest, or preference for complicated problems rather than those more easily mastered (p. 175-176).

The concept of evaluation goes far beyond simply grading. With growth as a concurrent concern, whether or not it is occurring would need to be ascertained. Chapman's work (1978) provides a plethora of ways for obtaining records of student work and learning which are subject to evaluation. These include portfolios of two dimensional work, photographs of three dimensional work, student diaries serving as records of responses to personal development, tape recordings of interviews and discussions, videotapes demonstrative of student involvement, and in class sign up sheets to chronicle responsible behaviors and activities. The interview and unobtrusive observation are cited as tools to more clearly ascertain how the student thinks, feels, and sees. Anecdotal records and checklists regarding attitudes and knowledge are seen as possible forerunners to periodic summary reports of student growth made jointly by student and teacher. Tests are presented as a manner of checking factual knowledge and the student's ability to recognize, interpret, and apply

concepts in a hypothetical situation.

Gaitskell, Hurwitz, and Day (Gaitskell, 1958; Gaitskell, Hurwitz, and Day, 1982) present evaluative strategies from a substantiveley different perspective. The influence of the behavioral objective and Bloom's Taxonomy (1956) is strongly in evidence as the authors link the pre-determined, desired educational outcome with "formal" evaluations including teacher written tests and standardized art tests. A few possible questions are cited. However, the examples cited appear to seek responses representative of learning of the "knowledge level" (p. 496) with no example items directed at the comprehension, application, analysis, synthesis, or evaluation levels (Bloom, 1956). The potential triviality of test items which are easy to compose was noted.

The authors give recognition to the challenge and complexity of evaluating developing expressiveness and artistic creativity within the student. Approaches such as checklists, anecdotes, portfolios, personal files, and narrative progress reports are subsumed under the heading of "informal" methods of evaluation.

Eisner's work, Educating Artistic Vision (1972) presented information about evaluating student learning which, by its very style, underscores the notion of the importance of the development of critical thinking skills by promoting their use by the reader. Evaluation, testing, and grading are defined with evaluation's mission seen as a process of securing information in order to improve the educational process. Testing is regarded as a method of sampling

student abilities; and grading, the assignment of a symbol indicative of the student's achievement relative to specified criteria. Eisner argued for the addition of brief evaluative statements with the parent/teacher conference to the grading process.

Eisner explored the relationship between the general view that an educator holds about evaluation and the view one holds about the educational process. Expressive objectives with unspecified outcomes in which valuable qualities within the work are discovered and displayed after the process necessitate a differing evaluative approach, classroom milieu, and level of risk taking than specified behavioral objectives. Eisner's approach runs parallel to Hull's (1984) in support of the development of the critical abilities of the student/learner with the end in mind of self critique and evaluation of progress.

Eisner's way of identifying characteristics of landmarks in creative activity, *vis a vis*, boundary pushing, inventing, boundary breaking, and aesthetic organizing, provides the reader with a basis for increased acuity in day to day recognition of creative activity. A variety of ways of collecting and recording information pertinent to the evaluative process are noted and have been expanded upon in subsequent work (Eisner, 1985a, 1985b).

Wilson's scholarly work (1971) on evaluation of learning in art education provided a massive statement on the complexity of harmonizing art education objectives with behaviorists' theories of evaluation. This work represents not only extensive research and a sensitive attunement to the actualities of art



education, but also the positive overlapping effect of his leadership role in the development of the art testing instrument of the National Assessment of Educational Progress (NAEP).

A hybrid taxonomy appears in Wilson's listing of art education behaviors. They include perception, knowledge, comprehension, analysis, evaluation, appreciation, and production. Test item examples with expanded descriptions of purpose for each of these classifications follows essays developing their meaning. Certain items given as examples of questioning approaches were developed by Wilson, but the majority were pulled from existing tests, both published and unpublished, ranging in date from 1940 with strong concentration in the late sixties. Wilson also includes support for the "fluid teaching structure" (1971, p. 556) frequently occurring in the art classroom for which behavioral objectives are not appropriate.

All of the aforementioned authors included summaries of stated goals of art education, philosophies and/or rationals, prior to their comments, essays, or chapters on the topic of evaluation of student learning. Each linked, to varying degrees the "means" (strategies, objectives, process) with the "ends" (disclosure, evidence of learning, grades) (Eisner, 1972). Each recognized and gave example of varying ways to gather information of student learning suitable for evaluative purposes, with Chapman providing the greatest diversity in techniques suitable for the classroom teacher and Wilson providing the most complex appraisal of the marriage of art learning and behaviorist measurement strategy through testing. Eisner provided the most verbal support for the

unspecified outcome with flexibility and artful teaching as keys to evaluation of this outcome. All writings support the notion that evaluation of student learning in the visual arts is grounded in a multiplicity of approaches.

### Proposed Curricular Changes and Surrounding Issues Related to Evaluation of Learner Achievement

The current directional focus within the field of art education calls for curricular reform including articulation of art disciplines as the basis for content in art curriculum. Clarification and explication of the kind and content of knowing/learning deemed characteristic of and of continuing value to the art education setting have inspired numerous essays. How one, as a leader/teacher, can show others that this process of knowing/learning is a fait accompli has brought open debate to the topic of evaluation of student learning in the visual arts. Hence, an exploration of essays and articles on curricular reform (primarily supportive to DBAE) and of commentary from art educators regarding the possible role of standardized testing is important to the study. Selections of literature were made with an eye to being au courant and date primarily from 1984 to 1988. However, one must note that actively defining the structure of the art education discipline has been in process for the past twenty five years, with DBAE as a label and specific focus for the past four years.

### Call for Curricular Reform and Current Directional Focus

The most direct route to an awareness of the curricular reform movement

in art education and its current directional focus may be by reading titles of recently published articles in a few art education and general education publications.

As with many reform movements, a critical point-in-time was the naming of the approach. Greer's 1984 article "Discipline-Based Art Education" provided the naming nuclei about which other scholars and authors began to cluster with supports, criticisms, and considerations. "Contemporary Reform and Contents of Curricula" (Packard, 1984), "Curriculum Validity in Art Education" (Broudy, 1985), "Art in Education: Five Rationales" (Schubert, 1986), and "Continuing the Translation: Further Delineation of the DBAE Format" (Diblasio, 1985) are but a sampling of the intensifying, exploratory work that was at hand in the mid eighties.

The concepts were not new, but rather more clearly defined. If art educators were to embrace these concepts, this named approach of DBAE with its quadripartitions of aesthetics, art criticism, art history, and art production, then further legitimating efforts in the name of the approach were to be expected. This came in the form of evaluation of student learning. "There is content that can be assessed and for which teachers can be accountable", stated the namer (Greer, 1984). "If art is to be taken as a serious discipline, instruction will have to exhibit characteristics that traditionally have distinguished instruction in the basics" commented the testing specialist (Hoepfner, 1984). "Evaluation is essential for validation of student achievement," echoed an art educator (Day, 1985). Such statements are common within the printed material which support

the DBAE approach.

The summer of 1987 edition of The Journal of Aesthetic Education was devoted entirely of DBAE with ten articles by art educators and was distributed free of charge to many within the field. The December 1987/January 1988 edition of Educational Leadership proliferated with articles about the status of art education. Included were interviews with Elliot Eisner on DBAE (Brandt, 1988a) and Howard Gardner on assessment in the arts (Brandt, 1988b) as well as reports of particular programming examples in public schools and their outcomes (Brickell, Jones, & Runyan, 1988; Greer & Silverman, 1988; Tollifson, 1988; and Wolf, 1988). Clearly we see a direction toward specified content and evaluation of student learning.

#### Commentary and Definition of Issues

Though the naming, shaping, and delineation of DBAE and its subsequent nameable, shaped, and delineated curricula have been and are being described, field tested, and funded (Burton, Lederman, & London, 1988; Greer & Silverman, 1988), these activities have not occurred in a vacuum. Basic controversial issues appear to have been 1) What is it we really want to include in our curricula? (Hausman, 1988; Hamblen, 1988; and Burton, Lederman, & London, 1988), 2) How are we going to protect the unspecified outcome, the student's expression? (Clark, Zimmerman, and Zurmeuhlen, 1987; and Hamblen, 1985 and 1986) and 3) How are we going to evaluate and disclose the outcomes that we do decide are important? (Clark, Zimmerman, &

Zurmeuhlen, 1987; Eisner, 1974; and Greer & Hoepfner, 1986).

With statements like "If we take instruction in art seriously and demand that expectations for student achievement be met then achievement tests are necessary" (Greer & Hoepfner, 1986, p. 43), simultaneously coupled with a lack of disclosure of the testing instruments which have been field tested in a DBAE setting (Hausman, 1988), it is no wonder that art educators are testy about testing.

The focus of this research has been state level visual arts assessment instruments including standardized tests. It is interesting to note that as art educators become more interested in the role of the standardized test, general educators, politicians, and parents are seriously questioning the role that the standardized assessment test has had in our public schools (Fiske, 1987; Shanker, 1988). As the impact that standardized assessment of student learning has had on educational diversity, curricula content, school autonomy, and student understanding and critical thinking is brought to debate in general education, art educators are considering which kinds of evaluative tools may most fairly and appropriately measure and disclose student learning.

#### Research in Evaluative Techniques in Art Education

Research in the evaluation and disclosure of student learning in the visual arts represented varied perspectives which ranged from holistic concerns of the effects of evaluative statements on students' artistic performance and motivation (Gerhart, 1986) to a technical concern of establishing reliability in

concept generalization testing (Clark, 1984). Institutions as well as individuals reported research in student evaluation.

The Educational Testing Service (ETS) had two major areas under consideration relating to assessment in the visual arts. One was the development of flexible modular tests to aid schools in assessing student progress and provide feedback regarding curriculum and instruction. The second was the development of materials for teacher inservice in evaluative techniques. (Sims-Gunzenhauser and Wanders, 1986).

In a pilot study in Pittsburgh, ETS, in conjunction with Harvard Project Zero, was also researching the usefulness of student journals and portfolios as possible innovative forms of assessment. Named ARTS PROPEL, parts of this research were presented at the NAEA conference in Boston in April, 1987, under the title of "Portfolios and Journals: The Integration of Making and Thinking in the Visual Arts" (Sims-Gunzenhauser and Wanders, 1986). When Howard Gardner was questioned as to why Harvard Project Zero was working with ETS and the Pittsburgh schools to develop ways to assess student achievement in the arts, he replied, "We want to know whether individuals involved in the arts are getting something out of their experiences, and whether those results can be documented" (Brandt, 1988. p.30).

Another group of researchers, Enid Zimmerman, Gilbert Clark, and Marilyn Zurmuehlen, presented a summary of research on art testing, methodology, content, and implications for the future at the spring 1987, NAEA conference ("Super Sessions," 1986). Subsequently a book has been

published by the National Art Education Association entitled Understanding Art Testing: Past Influences, Norman C. Meir's Contributions, Present Concerns, and Future Possibilities by Clark, Zimmerman, and Zurmuehlen.

In this book's concluding chapter, we find a revealing discussion by the three authors in which the issues of standardized testing, organic curriculum approaches, and the political realities of accountability for the visual arts are candidly addressed.

At the national level the earlier research and development which surrounded the National Assessment of Educational Progress (NAEP) in the Arts provided a plethora of information about test development and structuring of a visual art assessment. In no other instance has so much time and money been invested in a singular approach to evaluation of student learning in the visual arts as it was for the 1974-75 and 1978-79 sampling assessment by way of development, administration, scoring, and presentation of outcomes. Though independent in nature, Wilson's work (1971), "Evaluation of Learning in Art Education," could not help but have benefited from his exposure to and interactions with the numerous contributors to the NAEP as he fulfilled the role of chairperson at the national level.

At the states' levels, a variety of efforts have been made and continue to be made to form and use visual arts assesement instruments. Their work will be the primary topic of this study.

### Standardized Testing of Learner Achievement in the Visual Arts

General awareness of standardized testing of student achievement in the visual arts was essential in this study of state level visual arts assessment instruments. Three topic areas were identified as foundational to the study. These were 1) the presence of preexisting examples of standardized tests of learner achievement in the visual arts, 2) the development of questions which accommodate the unique characteristics of learning within the visual arts and their use within the format of the standardized test, and 3) the NAEP in Arts as a model for some states' assessment instruments.

#### Preexisting Examples of Standardized Tests of Learner Achievement in the Visual Arts

In search of preexisting examples of standardized tests of learner achievement in the visual arts, listings which would or could have included such tests were examined. These included Anastasi's "Classified List of Representative Tests" (1982, p. 669-682); Buros Institute of Mental Measurements' Tests In Print III (Mitchell, 1983); Clark's "Selected Chronology of Publications: Inquiry About Children's Drawing Abilities and Testing of Art Abilities" (Clark, Zimmerman, & Zurmuehlen, 1987); Wilson's references for the chapter "Evaluation of Learning in Art Education" (1971, p. 556-558); and the Educational Testing Service's "Annotated Bibliography of Tests, Art" (1987).

Anastasi noted in her chapter on educational testing that test may be loosely categorized on the basis of the measurement of developed abilities. "All



ability tests - whether they be designed as general intelligence tests, multiple aptitude batteries, special aptitude tests, or achievement tests - measure the level of development attained by the individual in one or more abilities." (1982, p. 394). Anastasi suggested that types of tests could be ordered along a continuum in terms of the degree of specificity of background experience that they presuppose. Course oriented and teacher made achievement tests would be considered to have the most specific information focus. Next would come broadly oriented achievement tests; then, verbal type intelligence and aptitude tests followed by non-language and performance tests; and, finally, "culture-fair" tests designed for use with persons of wideley varied backgrounds.

After review of lists of art tests with the thought of locating standardized visual arts achievement tests, it was found that tests existed at most of the points along the continuum. There were interest inventories, tests of creativity, occupational competency tests, attitude scales, cognitive assessments, nonverbal ability tests, diagnostic achievement tests, assessments of aesthetic judgement and perception, item banks, and measurable objectives pools. Tests had been developed by individuals, art educators, testing companies, school districts, universities, the national government, or educators from other countries.

Several tests appeared on more than one list. Notably absent were state level visual arts assessment tests.

### Development of Questions Suitable for Standardized Testing in the Visual Arts

The development of questions which focus on the unique characteristics of learning within the visual arts goes hand in hand with their inclusion within the format of the standardized test. At issue is suitability, not only of item type, but also of overall evaluative approach. This has been a long standing consideration in art education (Eisner, 1974).

Various art educators and test developers have explored question formulation and testing technique in areas specific to the DBAE movement in art education (Day, 1985; Greer & Hoepfner, 1986; Hoepfner, 1984). Also, the J. Paul Getty Trust in support of the discipline-based art education movement has employed scholars Hoepfner (1984) and Day (1985).

Though highly debated interest has been paid to the mingling of standardized testing and accountability with DBAE (Greer & Hoepfner, 1986; Hamblen, 1985, 1986, 1987; and Hodson, 1985) it may not be concluded that DBAE proponents see standardized testing as an exclusive evaluative tool. Day (1985) recommended consideration of student achievement by way of observations, interviews, discussions, performance, checklists, questionnaires, tests, essays, visual identification, attitude measurements, and aesthetic judgements.

However, item/question development in art education remains a challenge. It appears that, although listings of varied tests exist, a comprehensive pooling and exposition of possible items has not occurred and

been published since Wilson's work of 1971.

### NAEP as a Model of Test Development and Example for State's Assessment Instruments

The National Assessment of Educational Progress has twice conducted visual arts assessments at the national level of a sampling of public school students who were nine, thirteen, and seventeen years of age. The first testing was in 1974-75; the second, in 1978-79. Areas that were identified as having been assessed in the first testing were art production, knowledge about traditional Western forms of art, sex differences in artistic achievement, and the relationship of cognitive and affective objectives to art experiences. "Four areas on art learning were assessed: (a) Valuing Art, (b) Knowledge of Art History, (c) Responding to Art: Perceiving, Describing, Analyzing and Judging, and (d) Design and Drawing Skills" (Lovano-Kerr, 1985).

Numerous reports from the National Assessment for Educational Progress (NAEP) provided information regarding formulation of measurable objectives in art, test exercise development, scoring of performance items (drawing), and description of outcomes of the two assessments which were conducted (NAEP, 1977, 1978, 1981a, 1981b; Norris & Goodwin, 1971; Ward, 1982; and Wilson, 1975). These reports have impacted certain state level visual arts assessment testing. The Council of Chief State School Officers' report (1985) noted that:

"Such testing...is administered at traditional testing levels for academic

subjects (usually grades 4, 8, and 11). The most concentrated efforts have modeled assessment instruments after those designed for the National Assessment of Educational Progress. The increasing number of states actively pursuing statewide assessment programs indicates a definite trend toward standardized testing in the arts" (pp. 5-6).

NAEP's work provided an example of standardized visual arts assessment tests which have been field tested. The national assessment instruments have served as models for certain state level tests, have provided pools of testing exercises, and have served as a forum in which possible problems and limitations of standardized testing in the visual arts have surfaced.

### Criteria Applied to Standardized Tests

To become more familiar with the criteria applied to the evaluation of standardized tests as assessment instruments, it was necessary to 'cross over' into the test and measurements field. References were sought (American Psychological Association, 1974; Anastasi, 1982; Black & Broadfoot, 1982; Mehrens & Lehmann, 1980) which would help to clarify the expected traits, procedures, and roles of the standardized assessment instrument. Mehrens and Lehmann provided a succinct listing of traits of standardized achievement tests. These included: 1) specific instructions for standardized administration and scoring procedures; 2) test content determined by curriculum and subject matter experts involving extensive investigation of existing syllabi, textbooks, and programs done in such a manner so as to insure a fair sampling of content; 3) meticulous construction procedures that include constructing objectives and

test blueprints, employing item tryouts, item analysis, and item revisions; 4) in addition to local norms, standardized tests typically make available national, school district, and school building norms; and 5) the purposes and use of these instruments is best suited for measuring broader curriculum objectives and for interclass, school, and national comparisons 1969, p. 166).

Anastasi's "A Suggested Outline for Test Evaluation" was advisedly selected as a tool for review of the state level visual arts assessment instruments received in the course of this study which closely aligned themselves with criteria from the test and measurements field (Appendix D). Further explanation of Anastasi's criteria for test review and evaluation will follow in the design of the study.

### CHAPTER III

#### DESIGN OF THE STUDY

The study focused on reported state level visual arts assessment instruments, either in use, having been used, or in the process of development. It was determined that copies of existing tests, accompanying technical manuals, information regarding the manner in which the test was developed, or other publications about the visual arts assessment program would provide suitable data for analysis. Alternative techniques for the evaluation of student progress which were state- directed would also be noted.

#### Data to be Gathered

Twelve states and the District of Columbia were identified as having or developing visual arts assessments in the form of standardized tests or alternative techniques. This reporting of states appeared in the December 1986 issue of the NAEA (National Art Education Association) News in two ways: 1) Ten states were mentioned in an article about the report titled "Arts, Education and the States". The report had been prepared by the Council of Chief State School Officers and was based on replies to one question from a questionnaire.

The question read as follows:

"D. Testing: We would like to determine the kinds of examinations administered at the state level and the place of the arts in these tests. 1. Indicate each area and grade in which your state administers statewide, standardized, or other examinations by pacing the appropriate letter (use a, b, c or d) on the line below:

a=state developed competency-based

b=National standardized achievement

c=statewide assessment

d=other " (Council of Chief State School Officers, 1985, p.34).

2) Brief articles about six states' assessment programs also appeared in the same issue of NAEA NEWS. These were usually written by state level art specialists. In total twelve states, Connecticut, Delaware, Hawaii, Louisiana, Maryland, Michigan, Minnesota, Missouri, Pennsylvania, South Carolina, Utah, Wisconsin and the District of Columbia were noted as having some sort of visual arts assessment or information gathering of student involvement in the visual arts.

A listing of the twelve state visual arts specialists or the state superintendents of schools and their mailing addresses was compiled, in addition to the name and address of the superintendent of schools in Washington, D.C.

### Methods of Gathering Data

The method of gathering data consisted of three steps:

1) Letters were sent in May of 1987 to request copies of the tests used by the states, accompanying technical manuals, information about test development or test adoption, and publications about the visual arts assessment program

(Appendix A). The request for data included a letter of support for the study from Dr. Richard LaTour, visual arts education specialist for the state of Oregon.

(Appendix B.)

2) Follow up letters were sent in November of 1987 (Appendix C). The request for copies of tests and technical manuals was repeated. Also, a copy of the original request was included as well as a copy of Anastasi's "A Suggested Outline for Test Evaluation" (Appendix D).

3) If needed, follow up phone calls were made in January of 1988. It was also decided to include Indiana in the listing of states after an article about their development of a state level visual arts assessment test was printed in the December 1987 NAEA News.

### Analysis

It was anticipated that certain of the available assessment tests would not actually be involved with the development of visual arts assessments as reported or that certain states would have techniques of student evaluation other than standardized visual arts assessment tests. Certain state's instruments would lend themselves to description by use of Anastasi's "A Suggested Outline for Test Evaluation" (1982) which puts forth criteria considered to be important for standardized tests in the tests and measurement field. Data received in response to the letters of request fell into three categories and three techniques were employed to report the results.

A. The first category was composed of states whose replies



indicated that state level assessment in the visual arts did not exist or that visual arts assessment was included in fine arts assessments in an abbreviated fashion, perhaps three or four questions. These responses were described in a narrative style and include comments by respondents about the current status of visual arts assessment or approaches to evaluation of student learning in the visual art.

B. The second category consisted of responses which indicated substantial work in the development of techniques for evaluating student achievement in the visual arts at the state level. Due to the variety of kinds of work, either done in the past or in the process currently, these responses and their characteristic components were described by individual state in a narrative style.

C. The third category was that of states in which state level standardized visual arts assessment tests had been sent in response to the request for data. States were selected who sent visual arts testing instruments appearing to most closely approximate characteristics described as criteria for standardized testing instruments within the test and measurement field.

The technique employed to report the results included a listing of materials received, a description of the instrument within the framework of Anastasi's outline, and a summary evaluation. Anastasi's outline, "A Suggested Outline for Test Evaluation" (Appendix D) was used to review and describe the tests and their accompanying technical manuals (when available). Characteristics which were noted in accordance with Anastasi's outline

included: general information such as the title, author , publisher, time required to administer, and cost; a brief description of the purpose and nature of the test including the type of test, target population, nature of content, possible existence of subtests, and test items; practical evaluation featuring design of test booklet, editorial quality of content, appropriateness, ease of use, ease of administration, clarity of directions, scoring procedures, examiner qualifications and training, and face validity and examinee rapport. Technical information which was noted when available included norms, reliability, and validity. Reviewer commentary was sought.

## CHAPTER IV

### RESULTS

States which developed and used or were developing visual arts assessment tools were identified in this study as those reported by the NAEA News (December, 1986) and by the Chief State School Officers report entitled "Arts, Education and the States" (1985). There were twelve states which included Connecticut, Delaware, Hawaii, Louisiana, Maryland, Michigan, Minnesota, Missouri, Pennsylvania, South Carolina, Utah, and Wisconsin as well as the District of Columbia. Clarification of the existence of these testing instruments, their description with regard to testing standards, and inclusion of surfacing alternative evaluative techniques were factors which shaped the results of the study.

Responses varied, including statements to the effect that there were not tests in use, copies of actual tests, test specifications, technical manuals, results of tests, and forms of evaluation other than tests. A table was developed to record the type of response from the twelve states and the District of Columbia (See Table1).

The three categories of states presented in the results included: states

TABLE 1. Record of Responses to Requests for State Level Visual Arts Assessment Instruments and Supporting Documentation

State Contacted:	CT	DE	HI	LA	MD	MI	MN	MO	PA	SC	UT	WI	DC
Response	X	X	X	X	X	X	X	X	X	X	X	X	X
NO contact													
NO test		X		X	X			X		X		X	
Test sent	X						X		X		X		
Developing test							X				X		X
Evaluative tools			X										
Question Development						X							
Item specification						X	X				*		
Technical manual	X								X				
Results	X					X	**		X				

\* Test closely related to objectives in state level curriculum guide.  
Curriculum guide available.

\*\*Report of results in NAEA News article, December, 1986.

Indiana was included in the study after a planned state level visual arts assessment testing was reported in an NAEA News article, December, 1987

whose replies indicated visual arts assessment did not exist or existed in an abbreviated fashion; states whose replies indicated substantial research and development of visual arts questioning, tests, or alternative approaches; and states which sent standardized visual arts assessment tests.

States Whose Replies Indicated Visual Arts Assessment  
Did Not Exist or Existed in an Abbreviated Fashion

Delaware, Louisiana, Maryland, Missouri, South Carolina, and Wisconsin responded that they did not have state level visual arts assessment tests. Pennsylvania has visual arts learning assessment in the form of three or four questions in a fine arts/humanities assessment instrument.

Delaware

A telephone conversation with James Gervan, Delaware's Department of Public Instruction's Music and Art Specialist, on January 22, 1988, revealed that Delaware did not include the visual and performing arts in their state level assessment program. Mr. Gervan did indicate that effective teaching in general with the characteristic of effective evaluation of student learning was receiving increasing attention from the Department of Public Instruction.

Louisiana

Louisiana's response came from Myrtle Kerr, Supervisor for the Arts and Humanities at the state level in letter form dated May 12, 1987. Ms. Kerr stated

that the arts are not included in Louisiana's student testing though attempts have been made to have three or four questions pertaining to the arts included in testing instruments. The state economy was cited as a deterring factor in the testing of students, in general subject areas as well as in the arts.

Louisiana has completed an assessment of the state of the arts in the schools having questioned principals, teachers, arts specialists, parents, and students.

#### Maryland

James L. Tucker, Jr., Chief of the arts and Humanities for the Maryland State Department of Education stated in correspondence dated May 20, 1987, that Maryland does have a competency testing program, but assessment instruments for the visual arts have not been developed.

#### Missouri

From the State of Missouri's Department of Elementary and Secondary Education, Richard L. King, Coordinator of Curriculum Services replied that Missouri no longer administers or has available the statewide assessment instruments which were requested. Mr. King stated that there had been a component at both the sixth and twelfth grade levels in the assessment instruments which was called "Aesthetic Sensitivity". This section had not been limited to the visual arts but rather addressed the arts in general.

### South Carolina

Mac Arthur Goodwin, the State Art Consultant from South Carolina's Department of Education replied in a letter dated December 8, 1987, that South Carolina had not developed state level visual arts assessment test at that time.

### Wisconsin

In a telephone conference with Dr. Martin Rayala, the visual arts director for the Department of Public Instruction in Wisconsin on January, 8, 1988, it was learned that Wisconsin does not have state level testing in the visual arts. Dr. Rayala mentioned several aspects of evaluation of the visual art student's learning in Wisconsin: a state law requires assessment, an item bank exists through the state's Bureau of Assessment and Testing, local districts have primary responsibility for developing their evaluative systems, and, by September, 1989, local school districts are to have copies of their evaluative systems available upon request by state level officials.

### Pennsylvania

Material sent by Leann R. Miller, Educational Assessment Specialist for the division of Educational Testing and Evaluation of Pennsylvania's Department of Education was included in correspondence dated May 14, 1987. Ms. Miller sent copies of the items used in assessment instruments at five grade levels in 1986 which represented measurement of visual arts learning, an

interpretive manual for Pennsylvania's Educational Quality Assessment (EQA) program, and a copy of the 1986 Data manual which contained statewide norms and correlational data. Due to the limited number of questions about the visual arts (three or four questions per grade levels 4, 6, 7, 9, and 11 contained within a ten to twelve item arts and humanities section), it was determined that Pennsylvania's visual arts assessment existed in a very abbreviated fashion.

Within the arts and humanities testing sections, the number of questions which pertained to the visual arts varied among the grade level forms. Fourth and sixth grade had three test forms: some fourth graders had two questions out of ten about visual art, others had three questions out of ten. Some sixth graders had three visual art questions out of twelve arts and humanities items, others had four questions out of twelve. Grades seven and nine each had four arts and humanities test forms with a consistent three out of twelve visual arts' item count. However, the four forms at the eleventh grade level had two, two, three, and one visual arts' items respectively out of a total of twelve arts and humanities questions. Norming sample groups with varying numbers of items per domain would not appear possible, but reports indicated high internal consistency (Pennsylvania Department of Education, 1986a).

Another difficulty apparent in the comparability of test forms was the issue of item specification. Even when numbers of visual arts' test items within the arts and humanities sections were consistent, item focus varied. For example, at the seventh grade level the three items in Form A measure recall of 1) art forms, 2) architecture, and 3) art forms. The items of Form B measure recall of



1) colors, 2) colors, and 3) art forms. In Form C, items measure recall of 1) artist, 2) art terms, and 3) art terms. Form D items measure recall of 1) colors, 2) art forms, and 3) art forms. The specification 'art forms' referred to recall of media use and its end product, i.e. sculpture, papier-mache, ceramic process, and collage. The 'architecture' item required awareness of the definition of architecture. The 'color' items were definitions of warm, cool, and primary colors, and the 'art terms' sought an awareness of texture or calligraphy. It would seem that statistical norming, reliability studies, and content validity would be impossible without consistency of item specification among grade level forms. Also, monitoring sample groups' growth in a topic area over a period of time requires that subsets of questions remain constant from grade level to grade level.

It appeared that no visual arts' items at any grade level demanded student mental process beyond the recall level (Bloom, 1956; Stiggins, Rubel, and Quellmalz, 1986). No visuals were used in conjunction with items. No items elicited critical or evaluative response to the artwork on the part of the student, and no items assessed the process or production of student artwork.

Pennsylvania's EQA 1986 Data report appears to be a statistician's dream with copious tables of Pennsylvania school norms, percentile distribution of variables, correlation coefficients among school scores, correlation coefficients between school condition variable scores and school scores, and correlation coefficients among school condition variable scores and school

scores for each of five grade levels in eleven subjects. Subject areas assessed in 1986 were reading comprehension, writing skills, mathematics, analytical thinking, social studies, arts and humanities, science and technology, environment, health knowledge, self concept, and health practices. There was an apparent concern for a broad spectrum of student development. It would be of concern to art educators that there is such a limited numbers of items, inconsistency of item count and non-specification among grade level forms, as well as no questions which elicit higher level thinking, no questions which monitor students' capabilities of critiquing/ evaluating artwork, and no use of visuals, or a process/production component.

States Whose Replies Indicate Substantial Research and  
Development of Visual Arts Questioning,  
Tests, or Alternative Approaches

This categories of respondents included those states which have developed or are in the process of developing approaches to evaluation of learner achievement in the visual arts. Michigan developed and piloted, on a small scale, one of the first visual arts assessment instruments in the United States. Indiana and Washington, D.C. are currently in the test development and piloting process. Hawaii provided an example of an approach to evaluation of student achievement within a total art program evaluation model.

## Michigan

Information about Michigan's visual art testing came in the form of the "Analytical Report, Michigan Assessment Tests in Art, Statewide Sample 1977-78" sent on May 15, 1987, by Dr. Frank S. Philip, Fine Arts Specialist for the Department of Education. An accompanying letter from Dr. Philip stated that development of a new goals and objectives document was in process with completion projected for early 1988. Construction of new test items was planned.

Though the Analytical Report of 1977-78 did not reflect current status of visual art assessment within Michigan, it was evidence of substantial research and development of visual arts questioning in the form of a small scale pioneering state assessment. Michigan was the first state to conduct a statewide assessment of art learning. The report contained brief technical information (sample group demographics and test administration information), item samples with reporting of results for the 210 fourth and seventh graders tested; and, most interestingly, comments by Michigan art educators. Issues addressed by these art educators included items' construction, item interpretation by students as reflected in results, and visual arts testing in general.

Certain aspects of Michigan's assessment testing in art appeared to be common to state level art assessment development. The test content was linked to previously established state department of education performance

objectives. Consistency of test administration was achieved by having the same person give all tests for the sample within a particular school building. The sample selection was stratified through geographic area and community type (as was Connecticut's sample of their 1980-81 testing). A variety of item type was present. Some items were structured with slides or line drawings as visuals, and a production component was included. Scoring was done by an independent team. One of the scoring devices for production included a checklist of drawing strategies potentially used by students at the 3rd/4th grade level in the process of working through a production test item (similar in nature to a checklist drafted in Hawaii's evaluative model of 1987 as a student interviewer form which cites characteristics of student art work).

Difficulties were identified as having existed in either the structuring or outcomes of Michigan's testing instrument and process. It is unclear whether 3rd or 4th graders or 6th or 7th graders were tested. References were made to each grade level in the report. It is stated that the sample groups of fourteen students from each school were selected randomly with a procedure guaranteeing a cross section of talent and ability on the students' part. How this was achieved was not stated. The art educator's comments after reviewing results express concern for the limited amount of time (15 minutes) for the production test item. No question was raised about the ethics of posing a production experience requiring specific characteristics in the finished work which the student may or may not have within their visual repertoire. Also, exemplars were extensively used in the questions pertaining to knowing about

art with little comment about actual prior exposure for students or their economic or racial background. No women artists' work was used.

The Michigan Department of Education and the Michigan Art Education Association worked cooperatively on the preparation and evaluation of this assessment. The comments by art educators about the results and their interpretations which were included directly after numerical representation of item results contributed significantly to the final report. Considerable concern for growth and development of the curriculum and student was evidenced. Forthcoming recommendations included increasing students' critical skills and knowledge about art as well as overall strengthening of the curriculum.

The greatest historical significance of this document had to do with timing. The assessment was developed prior to the National Assessment for Educational Progress in the Arts publication of the Released Exercise Set in 1980. This Released Exercise Set of testing items/exercises impacted subsequent state level visual arts assessments markedly. Michigan's work clearly represents an independent pioneering effort representative of aspects, issues, and resolutions common to state level visual arts assessment development.

#### Indiana and Washington, D.C.

Both Indiana and Washington, D.C., have reported being in the process of developing and piloting visual arts tests. However, neither location has as yet sent copies of their testing instruments.

Telephone conversations, January 7th and 30th of 1988, with Jeffrey Patchen, Indiana's Fine Arts Specialist, and D. Cindy Borgman, who served as coordinator for the visual art assessment development, respectively, provided information of their pilot project.

Testing was planned for 65 fifth grade classrooms and 25 eighth grade classrooms. Teachers were to administer the tests and do the scoring. A pretest and posttest inservice was planned for all participating teachers. The tests were to have two sections: 1) a 65 item pencil and paper, multiple choice section with approximately one dozen color visuals and 2) a production section with motivation provided by a verbal scenario, 10 minutes of worktime, the showing of slides exemplifying the theme of the scenario, and 10 minutes more worktime. Scoring was to include student self evaluation regarding use of certain strategies within the artwork produced.

This assessment was classified as a diagnostic achievement test. Prior tests which were cited as contributing to the development of Indiana's test were the NAEP and tests from Minnesota. Indiana's tests were to be piloted in January and March of 1988.

Information from the District of Colombia came in a telephone conversation with Rena Watson, Washington's District Art Specialist. Ms. Watson stated that teachers had submitted materials and test questions commensurate with curricular content as part of the test development process. Tests were being piloted at the junior high and high school levels though they had not been officially adopted. A process of interactive decision making

between teachers at different levels appeared to be at work.

### Hawaii

Information regarding the assessment of visual arts learning in Hawaii came in the form of a paper titled "A Model for Evaluating Art Education Programs" by Morris K. Lai and Judy Shishido which had been presented at the American Educational Research Association annual meeting on April 20, 1987. Lai and Shishido's report boasted a healthy review of literature with mention of several art educators who have researched, developed, or supported approaches to the evaluation of art programs which may be considered holistic, qualitative, or naturalistic. The issue of combining qualitative with quantitative data gathering and analysis methods was addressed as well as the issue of suitability of standardized tests when local control of curriculum and the individual student's chance discovery are important (Lai & Shishido, 1987, p. 6). Lai and Shishido also explored the existence of standardized tests and reviews thereof which pointed to the lack of such tests and the reviewed tests' limitations. Research supporting teacher-made and criterion-referenced tests as resources for a test item pool was presented. Brief mention was made of discipline-based art education in as much as an art achievement test had been developed but not disclosed by the Getty Institute for Educators on the Visual Arts, and a recommendation for criterion-referenced testing had been forthcoming in a Getty report (p. 7). Expectancies that one may hold regarding children's artistic development were also outlined.

In consideration of these concerns, Hawaii developed an art program evaluation design. The first outlined step in the procedure was the targeting of grade levels 1, 3, 5, intermediate, and high school. It was stated that from "sub districts, classes from 2-7 elementary schools and 2 intermediate and high schools would be sampled. It is unclear how many classrooms or students this would total. Of this group a "20% stratified random sample subset" would undergo intense scrutiny (p. 9). The teachers of the sample classrooms, art specialists and art resource teachers, and state and district art education coordinators would respond to questionnaires which sought information about goals, how time was spent in the classroom setting, successes, art lesson assistance from the Department of Education, use of community resources, curricular priorities, material's supply, and training.

It appeared that from each sample classroom, the teacher would select three students, "a top student, an average student, and a bottom student", and would complete an evaluation of student performance form for each of these three. These performance forms were said to have four sheets, one each on the "four domains" (p. 10). Upon examination of a draft of the performance form, it was determined that the domains referred to were aesthetics, art criticism, art history, and production though specific reference to these categories was not made at any place in Lai and Shishido's report. Approximately 10-20 abilities were to be noted in each domain on a Remmer scale.

Teachers also would submit copies of tests they developed and sample work from each of the three students selected. The three students also would



complete a questionnaire which focused on students' judgements/evaluations, definitions, and attitudes toward and of artwork. These data gathering processes would occur in the sample classrooms.

In the classrooms to undergo intense scrutiny (20% of the aforementioned group of classrooms) the same activities would take place. However, the number of students selected for specific study would increase to nine, lesson plans would be submitted, and all students would do self evaluations. Teachers would be interviewed prior to three classroom observations by teams of three data gatherers. The team would include an observer taking field notes on preselected target students, a person operating a video camera, and an interviewer (connoisseur) with a 35 mm camera who would interview students and take slides of art work during at least one production session. The student interviewers would have guidesheets which would shape a recording of student expressive abilities, knowledge of artwork, knowledge of use of elements and principles of design, and recognition of aesthetic, cultural, and historical aspects of artwork. The interviewer forms would also provide a format for review of student artwork with 1) grade level estimation, 2) value judgement based on whether the student artwork was at, above, or below expected level, and 3) a checklist of characteristics which may possibly be present within student artwork at the actual grade level.

Non class program evaluation activities would include a review of Department of Education art resources, a review of resources listed by teachers in questionnaires, and a survey of people in the general population at a local art

museum to determine correlation between educational background and number of visits to visual arts displays annually.

The technique of the actual analysis of all of the data gathered in the above mentioned steps was not made explicit in the Lai & Shishido report.

Though the evaluation of student learning component within Hawaii's Model for Evaluating Art Education Programs represented a fraction of the whole program evaluation concept, certain significant assessment traits were identified. Production of student artwork in a non-test environment was given importance by way of slide making, video taping, interviewing and photography during student process, and checklists of artwork characteristics by interviewer and classroom teacher. Stratification of sample selection aided in validating the existence of individual differences. A variety of approaches toward recording and disclosing student experience and learning exemplified a holistic philosophy of evaluation.

#### States Which Sent Standardized Visual Arts Assessment Tests

Of the states who were reported to have visual arts assessments of student learning, three states responded with copies of standardized tests. These were Connecticut, Minnesota, and Utah. These tests appear to approximate established criteria for standardized testing instruments within the test and measurement field (Anastasi, 1982).

The techniques employed to report the responses within this section of

the study included a listing of materials received, a description of the instrument within the framework of Anastasi's Outline for Test Evaluation (Appendix D), and a summary evaluation. Not all topics noted on Anastasi's outline applied to the instruments received. In these cases, the topic heading was omitted.

### Connecticut

Information about the Connecticut Assessment of Educational Progress (CAEP) 1980-1981 Assessment of Art and Music was received in two mailings from Robert J. Saunders, Art Consultant, State of Connecticut Department of Education. The First mailing was received December 7, 1987, and included a brochure overviewing the CAEP Assessment of Art and Music and a "Summary and Interpretations Report" of CAEP Art and Music 1980-81. The second mailing of January 6, 1988, included a draft of the 1980-81 CAEP Art & Music Technical Supplement, a copy of the Grade 11 test, and instructions on "How to Score the Art Performance Items".

Connecticut's assessment represents the first state visual arts assessment using questions from the National Assessment for Educational Progress in Art Released Exercise Set of 1980). Connecticut's art assessment was also accompanied by extensive handling and reporting of statistical results.

Following are test characteristics which have been noted using Anastasi's Suggested Outline for Test Evaluation:

## A. General Information

### TITLE OF TEST

Connecticut Assessment of Educational Progress  
Art and Music 1980-1981

### AUTHOR(S)

A committee of art and music educators identified five goal areas for both music and art. The Art Assessment Advisory Committee had six members with Robert J. Saunders, Art Specialist, State Department of Education among them.

Goal areas for art were 1) to value art as an important realm of human experience, 2) to produce works of art, 3) to perceive and respond to aspects of art, 4) to know about art, and 5) to make and justify judgments about the aesthetic merit and quality of works of art.

Questions to match each goal area were chosen entirely from National Assessment of Educational Progress (NAEP) Released Exercise Set, 1980.

Dr. Brent Wilson, Chairman, National Art Assessment Committee, served as outside consultant.

### PUBLISHER, DATES OF PUBLICATION(S)

Questions were chosen from the National Assessment on Educational Progress in Art for 1974-1975.

Test was administered during 1980-81 school year.

"Summary and Interpretations Report" was prepared by National Evaluation Systems, Inc. in September 1981.

"1980-81 Connecticut Assessment of Educational Progress Art and Music Technical Supplement" October 1981

Time required to administer: 60 minutes

## B. Brief Description of Purpose and Nature of Test

### GENERAL TYPE OF TEST

Statewide visual arts assessment test administered to sample. Three domains measured: Cognitive domain consisted of multiple choice knowledge items. Performance domain required students to draw. Affective domain included questionnaire about student background and questions assessing student receptivity toward art.

"The tests contained questions to determine student attitudes and practices, as well as students' knowledge of basic art ... concepts. In addition, samples of students actually drew pictures ... in order to assess performance skills" (Connecticut State Department of Education, 1981).

### POPULATION FOR WHICH DESIGNED (age range, type of person)

Grades 4, 8, and 11

Tests were administered to intact (whole) public school classrooms.

Tests were administered to representative samples of Connecticut students at each of the three grade levels: about 3,500 fourth-graders in Winter 1981, 3,650 eighth-graders in Fall 1980, and 3,000 eleventh-graders in Spring 1981.

Two performance items, drawing exercises, were administered to a subsample of 5 in each class: 874 fourth-graders, 832 eighth-graders, and 720 eleventh-graders.

A Local Option phase of the assessment allowed local school districts to contract directly with National Evaluation Systems for services. This allowed local districts to 1) examine students, classes, schools, and districts as a whole, 2) examine special groups, 3) compare local achievement results with results gathered in the statewide testing as well as national results. Sixteen Connecticut school systems participated in the Local Option at one or more grade levels. Under this option, an additional 7,500 students were tested.

### NATURE OF CONTENT (e.g., verbal, numerical, spatial, motor)

Verbal (written) test items which were, at times, accompanied by visuals. Drawing performance items for selected sampling of students.

### SUBTESTS AND SEPARATE SCORES

Questions About You: family background, sex, age, art experiences, attitudes.

Section One: knowledge about art, making and justifying judgements about works of art.

Section Two: opinions, perceiving and responding to works of art.

Performance subsample: drawing.

Principal questionnaire: at the same time as the assessment was taking place, principals were questioned regarding art instruction within their schools, time spent, who teaches, problems, who schedules, and requirements.

### ITEM TYPES

Multiple choice, written.

Multiple choice, written, with prints or photographs as visuals.

Performance, drawing with specified outcomes.

## C. Practical Evaluation

QUALITATIVE FEATURES OF TEST MATERIALS (e.g., design of test booklet, editorial quality of content, ease of using, attractiveness, durability, appropriateness for examinees)

The test booklet which was sent was a xerox copy of the Grade 11 test. The visual design and layout were pleasing. Artwork was not identified as to title, artist, media, or ownership.

Items with answers requiring selection from multiple visuals on three pages would be confusing to the examinee.

### EASE OF ADMINISTRATION

No information available.

### CLARITY OF DIRECTIONS

Directions appeared clear on the 11th grade test.

### SCORING PROCEDURES

A bubble sheet answer form was used for all nonperformance items for the 11th grade test.

Performance items used NAEP scoring guidelines with an awarding of points for certain characteristics present in finished artwork.

### EXAMINER QUALIFICATIONS AND TRAINING

Examiners were local teachers who had been trained by district testing coordinators.

### FACE VALIDITY AND EXAMINEE RAPPORT

Face validity appears high in as much as the 11th grade test appears to be testing for goals of test developers.

Resulting average scores on cognitive items ranged from 40.4 to 53.5.

Examinee rapport may have been low due to unfamiliar material or questioning techniques of certain items.

## D. Technical Evaluation

### NORMS

On cognitive and performance items, the percent of correct responses per item was reported.

The affective results regarding student receptivity to art and student experiences were correlated with achievement on other items.

### STANDARDIZATION SAMPLE

A two-stage stratified cluster sampling design was used to select a random sample of students in grade levels 4, 8, and 11 based on actual distribution of student population in 1977-78 school year. Five types of communities (big cities, fringe cities, medium cities, small town/suburban, small town/rural) and six educational regions (Connecticut's educational service districts) provided the 2 variables of the sampling procedure. After procedures designed for random fairness to select of school testing sites, intact classrooms within these schools were randomly selected for test administration.

All students were considered eligible for testing except for those who were non-English-speaking or who were handicapped (physically, mentally, or emotionally) in such a way that they could not respond to

the test.

### RELIABILITY

Specific reliability procedures were not reported in summary and interpretations report.

The distribution of responses (in percentages) to each multiple choice cognitive item were reported along with goal being measured. There are types of reliability tests which can be performed with this information.

Scorer reliability on performance items was questioned because CAEP's scores were notably higher than NAEP's scores in this area. It was noted that CAEP trained local teachers to evaluate performance items whereas NAEP had a centralized staff of trained scorers.

Long-term stability of scores cannot be assessed with CAEP in Art because the test was administered only once.

### VALIDITY

Validation considerations may be inferred by attention to commentary about results and recommendations.

Content validity was brought to question in the Summary and Interpretations report with the statement that the "CAEP Test...by no means represents all of the important curriculum objectives" (p. 8).

Correlations were drawn between student achievement, attitude, and experiences.

#### E. Reviewer Comments

No reviewer comments were available either from Mental Measurements Yearbooks or Educational Testing Services resources.

#### F. Summary Evaluation

Major strengths of the Connecticut Assessment of Educational Progress in the Arts included: 1) Pioneered the use of items from the NAEP at the state level in a timely fashion. 2) Provided more data regarding the administration and scoring of performance items in settings which do not include professional scorers. 3) Exemplified the use of assessment professionals in test development and



technical reporting of results. 4) Provided base line data. 5) Generated recommendations for program improvements.

Weaknesses included: 1) Administered only once, measurement of changes in student growth, item redevelopment, and content alignment impossible.

### Minnesota

Information about the Minnesota Educational Assessment Program in the visual arts was received from William R. Allen, Assessment Specialist for the Minnesota Department of Education, in correspondence dated June 9, 1987. Mr. Allen sent two test booklets, an elementary and a secondary, which were dated 1982 and 1981 respectively. He also sent "Some Essential Learner Outcomes (SELO's) in Art" which formed the basis for the tests. Further data came from an article in the December 1986 issue of the NAEA News in which Mary Honetschlager, former Art Specialist for the State of Minnesota, reported on the two state art assessments which had been conducted. A telephone conversation with Dr. Raymond Higgins, current Art Specialist, on January 7, 1988, provided an update on the status of art assessment in Minnesota.

Minnesota's art assessment program represents a state visual art assessment using questions from the NAEP's Released Exercise Set of 1980 with no production component. Minnesota's art assessment was conducted twice. The extent of technical handling of data measured is unknown, as no technical manual was sent.

Following are test characteristics which have been noted using Anastasi's Suggested Outline for Test Evaluation and available information:

## A. General Information

### TITLE OF TEST

Minnesota Educational Assessment Program, Elementary Visual Arts, February, 1982.

Minnesota Educational Assessment Program, Secondary Visual Arts, September, 1981.

### AUTHOR(S)

Minnesota Department of Education, Division of Special Services, Assessment Section.

Primary Source of test items: National Assessment of Educational Progress (NAEP) in Art, Released Exercise Set (1980).

### PUBLISHER, DATES OF PUBLICATION

Some Essential Learner Outcomes (SELOS) in Art, primary basis for test item development and/or selection. Undated.

Visual Arts 4 Objectives, June, 1986; Visual Arts 8 Objectives, June, 1986; Visual Arts 11 Objectives, June, 1986. Listings of item specifications for tests. Items categorized as cognitive or affective.

Time required to administer: Approximately 45 minutes, a tape recorded reading of test items was used with timed pauses.

## B. Brief Description of Purpose and Nature of Test

Purposes of the tests were described as follows in an article by Mary Honetschlager, then Art Education Specialist with the Minnesota Department of Education, in the December 1986 copy of NAEA News.

Provide valid and reliable data on the knowledge, skills, and attitudes of Minnesota students.

Provide a means to evaluate curriculum strengths and needs.

Provide a means to focus developmental activities to improve student learning.

Measure change in student performance over time.

Provide materials, data, technical assistance to local school districts engaging in curriculum evaluation through the Local Assessment

Option.

### GENERAL TYPE OF TEST

Satewide visual arts assessment test administered to sample. Two domains measured: Cognitive domain consisted of multiple choice knowledge items. Affective domain included questions about student background and questions assessing student receptivity toward art.

### POPULATION FOR WHICH DESIGNED(age range, type of person)

4th, 8th, and 11th graders in Minnesota public schools.

### NATURE OF CONTENT (e.g., verbal, numerical, spatial, motor)

Written test items at times accompanied by visuals.  
No performance component.

### SUBTESTS AND SEPARATE SCORES

Item specification in SELO's identifies two possible subcategories: cognitive and affective.

Correlation of student attitude and experience with achievement would have been possible but information received does not indicate whether or not this process was pursued.

### ITEM TYPES

Multiple choice, written.  
Multiple choice, written, with prints or photographs as visuals.

## C. Practical Evaluation

### QUALITATIVE FEATURES OF TEST MATERIALS

Test booklets were well assembled. Prints or photographs were fairly clear, black and white. One print of a painting in each package of the elementary test was in color. Two prints of paintings in each package of the secondary test were in color.

### EASE OF ADMINISTRATION

A paced tape recording which read each item was used with the test booklets.

### CLARITY OF DIRECTIONS

Written directions appeared clear.

### SCORING PROCEDURES

A bubble sheet answer form was used by examinees.

### EXAMINER QUALIFICATIONS AND TRAINING

Unknown. It may be assumed that bubble sheet answer forms were computer read.

### FACE VALIDITY AND EXAMINEE RAPPORT

Face validity appeared high in as much as the elementary and secondary tests reflected item specifications in the Visual Art Objectives.

Reports on actual outcomes were not available which made inferences about examinee rapport inappropriate.

## D. Technical Evaluation

### NORMS

Not available.

Norms appear to have been tabulated in overall percentage of correct responses on, because comparisons of 1981-82 results to 1985-86 results was reported in terms of percent increase or decrease of grade level scores on specified objectives.

### STANDARDIZATION SAMPLE

5% random sampling of 4th, 8th, and 11th graders across the state.

### RELIABILITY

No technical manual available.

A parallel form reliability procedure may have been used as test booklets contain two similar forms at each level.

Equivalence of forms appeared constant with the exception that in the secondary test Package 1 had 33 cognitive items and 45 affective items and Package 2 had 30 cognitive items and 57 affective items.

Long-term stability was impossible to determine as the conditions

within Minnesota's public schools changed between the two testings due to a substantial reduction in art teaching staff, and the testing instruments, visual arts objectives, and SELO's are in the process of revision.

### VALIDITY

No technical manual.

#### E. Reviewer Comments

No reviewer comments were available either from Mental Measurements Yearbooks or Educational Testing Services resources.

#### F. Summary Evaluation

Major strengths of the Minnesota Educational Assessment Program in the visual arts include: 1) Minnesota as sole state with repeated use of testing instrument which used items from the NAEP in the Arts. 2) Recommendations generated for improvements based on test results were widely distributed and used to instigate change at local, state, and legislative levels. 3) Test results impacted course content with more emphasis on art criticism and art history. 4) Assessment impacted acceptance of art as basic subject. 5) Data used as evidence for need of improvement in art education.

Weaknesses included: 1) Lack of performance items. 2) Possible difficulties with technical procedures unknown to this study due to non-availability of technical manual.

### UTAH

Two sets of tests titled art inventories were received from Utah in response to requests for state level visual arts assessment instruments and supporting documentation. The first set, copywritten 1985, was sent by Charles B. Stubbs, State Art Director, on May 14, 1987. The second set, copywritten 1986, was sent by Ellis C. Worthen, Granite School District Art and Music Director on November 18, 1987. No technical manuals were sent. Further

information was gathered in a telephone conversation with Dr. Darrell Allington, retired Granite School District Art Specialist, on January 14, 1988.

Utah's art testing instruments represented the work of the Granite School District and the Institute for Behavioral Research in Creativity under the auspices of the Utah State Department of Education. Though not administered throughout the state at the time information was gathered, it was reported that plans were to have state wide use of tests. Utah has a state level art curriculum guide with sequential, specified outcomes, and these tests have been developed with alignment to curriculum in mind.

Following are test characteristics which have been noted using Anastasi's Suggested Outline for Test Evaluation:

#### A. General Information

##### TITLE OF TEST (including editions and forms if applicable)

Art Inventory: Test A; Art Inventory: Test B - 1985  
 Art Inventory: Test C; Art Inventory: Test D - 1985  
 Art Inventory: Test E; Art Inventory: Test F - 1985

Art Inventory: Test A, Forms I and II - 1986  
 Art Inventory: Test B, Forms I and II - 1986  
 Art Inventory: Test C, Forms I and II - 1986

##### AUTHOR(S)

Developed by Granite School District and the Institute for Behavioral Research in Creativity, Robert L. Ellison, Director, under a Title IVC Developmental Grant for Improvement of Art Instruction.

##### PUBLISHER, DATES OF PUBLICATION

Copyrights 1985 & 1986  
 Utah State Board of Education  
 Salt Lake City, Utah

Time required to administer: not known.

No technical manual was sent.

The following comments will focus on the 1986 versions of the Art Inventory Tests. Though most items are the same or similar to the items of the 1985 versions, the organization has changed.

#### B. Brief Description of Purpose and Nature of Test

Stated purpose: The purpose of this test is to find out what students have learned about art and how they like it.

#### GENERAL TYPE OF TEST

Tests appear to measure individual achievement in knowledge of selected elements and principles of art, ways of achieving perspective, color blending results, correctness of artwork, and knowledge of historical exemplars or recognizing types of exemplars.

#### POPULATION FOR WHICH DESIGNED

Public school student, K-12, Granite School District, in process of expanding to Utah public schools in general.

Test instructions in the three 1986 grade levels (A, B, & C) include items for student to note grade level, age, and sex.

Test A appears to be K-6; Test B, middle school or junior high; and Test C, high school.

#### NATURE OF CONTENT

Written (verbal) items and written items with visuals.

#### SUBTESTS AND SEPARATE SCORES

Not known, no technical manual.

Results for particular groups of items could be reported as subtest, i.e. items referring to knowledge of color.

### ITEM TYPES

Tests A, B, and C use three item types:

Tests A and B use 1) verbal multiple choice with black line visuals and 2) verbal multiple choice.

Test C uses 1) verbal multiple choice with black line visuals, 2) verbal multiple choice, and 3) verbal multiple choice with black and white prints of exemplars.

### C. Practical Evaluation

#### QUALITATIVE FEATURES OF TEST MATERIALS

Content appears to measure knowledge about art, primarily vocabulary.

No items ask for attitude or experience of the student though one stated purpose of the test is finding out how students like art.

Artwork used as exemplars was not identified as to title, artist, media, or ownership.

Test cover sheet is traditional, does not appear to be 'art' test, no visuals on cover.

Black line visuals are clear; black and white prints of exemplars which are used in Test C, Forms I and II, are small with obscured details.

Appropriateness of having only one test for K-6 is in question.

#### EASE OF ADMINISTRATION

All tests are read aloud and students are instructed to "Read the questions to yourself while they are being read aloud."

#### CLARITY OF DIRECTIONS

Directions appear clear though ambiguity of certain test items may confuse the discerning examinee.



### SCORING PROCEDURES

Test A, K-6, answers are marked in test booklet.  
Test B & C, 7-12, a separate answer sheet is used.

Actual scoring procedures not known, no technical manual available.  
However, all items appear to have one specific right answer.

### EXAMINER QUALIFICATIONS AND TRAINING

Classroom teachers serve as examiners. Tests are in connection with curriculum guide, inservice training has been provided in Granite School District.

### FACE VALIDITY AND EXAMINEE RAPPORT

Tests do not measure whether the student likes art as specified in test purpose.

Test A, Forms I and II, K-6, lower primary student and upper intermediate student may be disturbed by relative difficulty or simplicity of vocabulary.

Test C, Form I and II, high school, examinee may be disturbed by difficulty in seeing exemplars.

## D. Technical Evaluation

### NORMS

No technical manual available, though it appear that the Institute for Behavioral Research in Creativity (IBRIC) was processing a variety of resulting data (i.e. Dr. Darrell Allington mentioned statistical processes used in item analysis for test development).

### STANDARDIZATION SAMPLE

Unknown

### RELIABILITY

Unknown

Parallel forms of tests indicate the process or potential process of test reliability study.

Classroom teacher may be scorer and apparently has access to results in order to analyze effectiveness of teaching.

Equivalence of forms appears constant. Forms I and II of Tests A,B, and C had paralleling formats, content, and number of items.

Certain items are used in all test levels, A, B, and C.

### VALIDITY

No technical manual.

According to Dr. Allington a point bi serial formula was being used to determine item validity with  $+ .30$  as a standard.

### E. Reviewer Comments

No reviews were available from Mental Measurements Yearbooks or Educational Testing Services resources; however, critical commentary has appeared in the Journal of Curriculum Studies (Bullough & Goldstein, 1984), Studies in Art Education, Design for Arts in Education, and Art Education (Hamblen, 1985, 1986, and 1988). A major criticism was that the Utah curriculum, Art Is Elementary, and accompanying testing is representative of "technocratic mindedness" which reduces "moral, aesthetic, educational and political issues to technical problems" (Bullough & Goldstein, 1984, p. 144). Concerns are voiced by both Bullough and Goldstein and Hamblen for the maintenance of the position of the self-expressive, idiosyncratic, and serendipitous in teaching and learning when specified, testable outcomes are superimposed on the art class setting. The issue of the test and the testable leading the curriculum is raised, as well as the overwhelming presence of curricular objectives leading to application of skills to the detriment of objectives fostering expression, exploration, revelation, or creativity. According to the aforementioned authors, the major shortfalls of the testing instruments are intrinsically linked to shortcomings within the curriculum, Art Is Elementary, and the desire of its developers to control and have art be like other subjects.

### F. Summary Evaluation

Given: 1)that the Utah Art Inventories are course oriented achievement tests and not statewide visual arts assessment instruments and 2)that no technical manual was available.

Major strengths of Utah's Art Inventories include: 1) Tests are part of a package which has been created with the expressed motivation of maintaining the importance of art within the curriculum. 2) Use of item analysis in the test development process. 3) Variety of item structure. 4) Repeated field testing.

Weaknesses include: 1) The practice of testing all students on a regular yearly basis rather than sampling on a three to four year cycle may manifest an omnipresent limitation on the teaching and learning environment. 2) The tight alignment to curriculum guide and subsequent narrowness of content may restrict spontaneity and expressive outcomes in the art of teaching and learning. 3) There are no production or affective items, no measurement of student attitude and experience. 4) There are items which art educators may find inappropriate, incorrect, or insignificant for art learning.

## CHAPTER V

### DISCUSSION

This study's concluding commentary includes and branches from the issues of state level visual arts assessment instruments to the broader issue of evaluation of student learning in the visual arts. The content of the discussion will include a brief summation of the resulting answers to the research questions, recommendations for further research, surfacing implications for art education, implications for general education, and closing comments. Hopefully, a primary outcome of the discussion section will be the reader's recognition of complexity of the issue of evaluation of student learning in the visual arts.

In reviewing and summarizing the answers to the research questions, the following conclusions were made:

1. What form have the state visual arts tests taken?

The twelve states and the District of Columbia which were contacted in the process of information gathering for this study reportedly had state level visual art assessment programs and possibly tests. As one can see from the results, the forms ranged from non-existent to holistic to traditional assessment

testing instruments with a slant toward representation of art education values.

2. What do these instruments measure?

The visual arts assessment instruments in the form of tests provide a sampling of student learning. They are not comprehensive nor do they measure or disclose many aspects of student learning in the art classroom. They do measure and disclose certain specified cognitive outcomes, certain affective responses and attitudes, and certain production characteristics in a delineated and timed setting.

3. Does it appear that state level curriculum guides or learning objectives in the visual arts were taken into consideration in test construction?

Yes, generally speaking it appeared that a thoughtful evolutionary process of development of viable assessment instruments had been planned. Leaders considered not only the classroom teacher's contributions and ideas but also existing state level curriculum or the development and delineation of that curriculum. In some cases university level art educators were included in the planning process, in other cases available university resource people were notably absent. Despite the mention by state leadership of encountered funding "walls" for research and development, administration, and scoring of evaluative tools which would adequately reflect art education values and objectives set forth by assessment development committees, the leadership evidenced substantial efforts to sustain and clarify these values and objectives.

4. Did the tests' characteristics correspond to characteristics considered to be important by experts in the test and measurement field?

The actual testing instruments did utilize techniques and practices valued by the test and measurements field. Though at times the use of these techniques and practices appeared clumsy and naive, at times their use was extraordinarily sophisticated and superimposed. Perhaps a better question here would be, did the tests' characteristics correspond to characteristics considered to be important in the art education field?

5. What alternative techniques for evaluating student progress in the visual arts surfaced at the state level?

The holistic evaluative approaches to the student learning experience evidenced in Hawaii's work served as the most extensive example of alternative techniques to surface within this study.

### Recommendations for Further Research

Research cannot help but spawn newly formed research questions. In order to promote valid, ethical representation of student learning and growth, continued research in measurement, chronicling, and disclosure of the art student's learning is necessary. The following brief proposals for research were deemed important to the clarification of the issue of assessment of student learning in the visual arts. They reflect concerns for wise use of existing resources, for student rights to expression, for clarification of assessment standards valued by art educators, for fair treatment of individuals in a pluralistic society, and for field testing of varied approaches to evaluation of student learning.

The first proposal is for research resulting in the pooling and exposition of item

types on existing tests. What do these items disclose that is valued by art educators in student learning? We have disparate evaluative instruments which have sprung forth fully titled from varied, capable sources. There is no up-to-date comprehensive item bank for art educators which sorts possible item constructs as to suitability for grade levels or content specifications.

The second proposal is for research which would continue exploration into the effects of evaluative techniques on student motivation and commitment. Those of us that have felt and tasted in the air that change that can take place in student attitude after a grading/evaluative process recognize the need for sensitivity toward and empowerment of the student's expressions and purposefulness. Further research could support justification and advocacy for the student's rights to self and shared evaluation in settings responsive to student expression.

The third proposal is for the development of a screening device for critiquing proposed assessment procedures which represents art education values. Current devices such as Anastasi's "Suggested Outline for Test Evaluation" were constructed with other disciplines and traditional test and measurement approaches in mind. Though numerous considerations from the test and measurement field are important and useful in assessment of the art student's learning, tests have been noted as sampling student learning with students being fractionally represented when numerical results alone are used as evidence that learning has taken place (Finlayson, 1988, p. 5). A screening device for art educators to use in critiquing evaluative models would more sensitively represent art education values and whole student learning.

The fourth proposal is for research into the uncovering and elimination of bias in

art testing instruments. Are individual differences respected? Do evaluative instruments reflect considerations of student background and expression in a pluralistic society?

The fifth proposal is for continued field testing of multiple approaches toward the evaluation of student learning. An awareness of the input and experience of the classroom teacher is imperative. What works? What ideas do K-12 teachers have about evaluation of art students' learning. How much time do art teachers have for the process? What are the practical considerations?

### Implications for Art Education

Implications of the study for the field of art education may best be exemplified by recognition of three issues. These issues are the role of evaluation in the evaluation setting as it relates to current curricular changes, the need for dialogue, and the need for a clear recognition that differences exists between specified and unspecified outcomes.

The role of evaluation of student learning in any educational process demands recognition. Traditionally 'testing' has lead the curriculum. With DBAE and subsequent efforts to articulate, sequence, and specify learning outcomes, art education has entered a tenuous time. We are in the position of simultaneously resolving curricular and evaluative issues. We can look to general educators and examine strengths and weaknesses. We can look to art educators and find an amazing pool of independent, innovative, responsive, and capable professionals. Never before has a setting existed in art education in which the tools of orchestration of



multiple purposes were more needed at the national, state, and local levels. These very tools are commonly used within the art learning setting. Fortunately, we, as art educators, represent our own best resources.

The need for dialogue on quantitative versus qualitative issues regarding evaluation of the art student's learning and its disclosure exists. How can art educators in the public school classrooms, universities, and state administrations communicate with one another? Pooling of information in a conference setting which facilitates communication, synthesis, mutual respect, and balance followed by the timely survey of opinion from the field may serve to facilitate dialogue.

A clear recognition of the roles of the specified, measurable, testable objective and the unspecified, unique, expressive objective within the art learning setting needs to be established. The need for recognition of student experience in a field in which ideosyncratic development is seen as essential, embracing "right answer" ideals without subjecting them to close scrutiny and modification may be not only inappropriate but also unethical. Clarity of the different roles of objectives lends itself to advocacy of student and teacher's rights to ideosyncratic development.

### Implications for General Education

Research and development of evaluative devices in the field of art education may mean significant contributions to the field of general education. Contributions could include serious questioning of the test as an accurate representation of student learning, advancement in the development and monitoring of critical thinking skills, and student and teacher empowerment.

The viewpoint of the test as a sampling device provides the potential for the arts to lead in the use and support of multiple approaches to the evaluation of student learning. Possible outcomes include a rejection of the notion that fractional, numerical representations of student learning are in the best interests of education.

The critical thinking skills movement and art education's critical components are mutually supportive. Astute questioning and self questioning leading to positioning a critical stance have long been traditions within the art education field. Developing ways to further promote, monitor, and disclose this process are currently at hand and have the potential for providing an artistic overlay to impact the critical thinking skills movement.

Student and teacher empowerment are the natural outgrowths of the development and fostering of critical thinking skills. In order to avoid the paradox of developing critical thinking skills within the individual and then disallowing the individual's rights to exercise those skills in the process of evaluation of their learning/teaching, one must promote ownership of the evaluative process by the individual in question. Guided self-evaluations and disclosures can promote independence and strengths.

### Closure

What we as art educators value in evaluation of student learning will indeed lead our curricula and our field. We have before us a time of definition, change, and directedness as we reformulate curriculum while simultaneously

redefining assessment. What a more perfect time to put forth our ideals.

Recognition of general education's disenchantment with "back to basics" and standardized tests may provide the impetus for the field of art education to help us avoid anachronistically embracing restrictive evaluative instruments. Careful concern for and structuring of evaluative instruments can serve to strengthen rather than compromise art education's role within the general education setting in this adaptation process which serves accountability.

Astute concern for the whole individual as learner may lead to a new rationale for art education in the information age. As we promote within the individual those abilities to critique, screen, and compose information through artful self-construction of limits and uses, we foster balance.

Individual development that promotes the ability to see the whole system and filter the irrelevant is imperative in the dawning information age. Drawing the line is needed for survival. Drawing an artful line is needed for peace of mind.

APPENDIX A  
LETTER OF REQUEST  
FOR DATA



## UNIVERSITY OF OREGON

Dear

It is my understanding that your state administers a visual art assessment test, possibly at three different grade levels, as a measurement of student progress in the visual arts.

I am currently conducting research as part of my master's study at the University of Oregon in the Department of Art Education with regard to the content of these state level visual arts assessment tests.

Would you be able to send me copies of the tests that are used by your state, as well as any accompanying technical manuals? Additional information regarding the manner in which the test was developed or adopted or publications about your visual arts assessment program would be most appreciated. Also, could you give me the name and address of a particular person to whom I should address further questions?

The purpose of this study will be to do a content analysis of the tests, or selected tests, in order to see what content areas within the field of art actually are being tested at the state level at this time.

I have identified twelve states (Connecticut, Michigan, Minnesota, Pennsylvania, Utah, Missouri, Delaware, Hawaii, Louisiana, Maryland, South Carolina, and Wisconsin) and the District of Columbia as currently having a visual arts assessment process.

The limited number of states administering visual arts assessment tests means that every potential respondent is important to my study.

I would be happy to share the results of the study in the form of a summary report should you indicate an interest. Hopefully, the information derived will be helpful to participating states who

may wish to update their testing instruments, states who are in the process of development of testing instruments, and art educators who wish to update their knowledge of current expectations within the field.

This research is being conducted under the direction of Dr. Linda F. Ettinger, Assistant Professor, Department of Art Education, University of Oregon, Eugene, Oregon. Please feel free to call either of us at (503) 686-3639 if you have questions.

I look forward to hearing from you and receiving test samples and supportive materials from your state.

Thank you for your time.

Sincerely,

Sandra Finlayson  
Graduate Teaching Fellow  
Department of Art Education

SF:cw

APPENDIX B

LETTER OF SUPPORT FROM

RICHARD LATOUR



OREGON DEPARTMENT OF EDUCATION  
700 PRINGLE PARKWAY SE, SALEM, OREGON 97310-0290 PHONE (503) 378-3569

May 1, 1987

Dear Art Education Colleague:

I am writing to urge your support and participation in this study of visual arts assessment instruments. At the national, state, and local levels, we are advocating quality art education programs. Achieving that quality, in part, requires an understanding of the alignment between our goals, our instruction, and our assessment instruments and practices. As we move forward with our improvement efforts, the need and benefit of research on assessment will be critical to our success. This study represents an important component for increasing our knowledge base in evaluation as it relates to art education. I hope you will consider your participation as both important and beneficial to our field.

Cordially,

A handwritten signature in dark ink, appearing to read "Richard M. LaTour".

Richard M. LaTour, Ph. D.  
Visual Arts Education Specialist  
Oregon Department of Education  
(503) 378-3602

ey86436



APPENDIX C

SECOND LETTER OF REQUEST

FOR DATA



## UNIVERSITY OF OREGON

Sandra Finlayson  
Room 252, Lawrence Hall  
Department of Art Education  
School of Architecture & Allied Arts  
University of Oregon  
Eugene, OR 97403

November, <sup>12</sup> 1987

Dear

Last spring I sent a letter to you requesting copies of your state level visual arts assessment tests and their accompanying technical manuals. You did not respond, neither by sending the requested materials nor by telling me that you do not have assessments in the visual arts. As I start to process the information from the states which did respond, I find myself to be curious still about your programs.

Would you have time now to respond to my inquiry? Copies of the original letters requesting information have been enclosed.

If you do have state level visual arts assessment tests and their technical manuals and will share the information, it could be beneficial to my study and to the study of standardized evaluation of student achievement within the visual arts field. We, art educators, do have many questions and concerns about appropriate forms of student evaluation and need to be familiar with current developments.

I will be presenting my findings to the art education community and would like to have an accurate representation of the state level tests. Your response and support would be greatly appreciated.

Sincerely,

Sandra Finlayson

Enclosures

APPENDIX D

ANASTASI'S "A SUGGESTED OUTLINE  
FOR TEST EVALUATION"

Anastasi, A. (1982). Psychological Testing (5th ed.) New York: MacMillan. p. 666-667.

## A Suggested Outline for Test Evaluation

### A. General Information

Title of test (including editions and forms if applicable)

Author(s)

Publisher, dates of publication, including dates of manuals, norms, and supplementary materials (especially important for tests whose context or norms may be outdated)

Time required to administer

Cost (booklets, answer sheets, other test materials, available scoring services)

### B. Brief Description of Purpose and Nature of Test

General type of test ( e.g., individual or group, performance, multiple aptitude battery, interest inventory)

Population for which designed (age range, type of person)

Nature of content (e.g., verbal, numerical, spatial, motor)

Subtests and separate scores

Item types

### C. Practical Evaluation

Qualitative features of test materials (e.g., design of test booklet, editorial quality of content, ease of using, attractiveness, durability, appropriateness for examinees)

Ease of administration

Clarity of directions

Scoring procedures

Examiner qualifications and training

Face validity and examinee rapport

#### D. Technical Evaluation

##### 1. Norms

Type (e.g., percentiles, standard scores)

Standardization sample: nature, size, representativeness, procedures followed in obtaining sample, availability of subgroup norms (e.g. age, sex, education, occupation, region)

##### 2. Reliability

Types and procedure (e.g., retest, parallel-form, split-half, Kuder-Richardson or coefficient alpha), including size and nature of samples employed

Scorer reliability if applicable

Equivalence of forms

Long-term stability when available

##### 3. Validity

Appropriate types of validation procedures (content, criterion-related predictive or concurrent, construct)

Specific procedures followed in assessing validity and results obtained

Size and nature of samples employed

E. Reviewer Comments

From Mental Measurements Yearbooks and other sources

F. Summary Evaluation

Major strengths and weaknesses of the test, cutting across all parts of the outline

## BIBLIOGRAPHY

American Psychological Association. (1974). Standards for educational and psychological tests. Washington, DC: Author.

Anastasi, A. (1982). Psychological testing. New York: Macmillan.

Black, H., & Broadfoot, P. (1982). Keeping track of teaching assessment in the modern classroom. Boston, MA: Routledge & Keyan Paul.

Bloom, B. S. (Ed.) (1956). Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain. New York: David McKay.

Brandt, R. (1988a). On assessment in the arts: A conversation with Howard Gardner. Educational Leadership, 45(4), 30-34.

Brandt, R. (1988b). On discipline-based art education: A conversation with Elliot Eisner. Educational Leadership, 45(4), 6-9.

Brickell, E. E., Jones, N. T., & Runyan, S. H. (1988). An art curriculum for all students. Educational Leadership, 45(4), 15-16.

Broudy, H. S. (1985). Curriculum validity in art education. Studies in Art Education, 26(4), 212-215.

Bullough, R. V. & Goldstein, S. L. (1984). Technical curriculum form and American elementary school art education. Journal of Curriculum Studies, 16(2), 143-145.

Burton J., Lederman, A., & London, P. (Eds.). (1988). Beyond DBAE: The case for multiple visions of art education. University Council on Art Education.



- Chapman, L. (1978). Evaluating children's progress in art. Approaches to art in education. New York: Harcourt Brace Jovanovich.
- Clark, G., Zimmerman, E., & Zurmuehlen, M. (1987). Understanding art testing. Reston, Virginia: National Art Education Association.
- Clark, G. A. (1984). Establishing reliability of a newly designed visual concept generalization test in the visual arts. Visual Arts Research, 10(2), 73-78.
- Clark, G. A., Day, M. D., & Greer, W.D. (1987). Discipline-based art education: Becoming students of art. Journal of Aesthetic Education, 21(2), 129-193.
- Connecticut State Department of Education. (1980). 1980-81. Connecticut assessment of educational progress. Art. Grade 11 (Test). Hartford, Connecticut: Author.
- Connecticut State Department of Education. (1981). CAEP. 1980-1981 Assessment of art and music. Hartford, Connecticut: Author.
- Cornia, I. E., Stubbs, C. B., & Winters, N. B. (1976). Art is elementary: Teaching visual thinking through art concepts. Provo, Utah: Brigham Young University Press.
- Council of Chief State School Officers. (1985). Arts education and the states. a survey of state education policies. Washington, DC: Author.
- Day, M. D. (1985). Evaluating student achievement in discipline-based art programs. Studies in Art Education, 26(4), 232-240.
- Diblasio, M. K. (1985). Continuing the translation: Further delineation of the dbae format. Studies in Art Education, 26(4), 197-205.
- Educational Testing Service. (1987). ETS test collection. Annotated bibliography of tests. Art. Princeton, New Jersey: Author.

- Eisner, E. W. (1972). Children's growth in art: Can it be evaluated? Educating artistic vision. New York: Macmillan.
- Eisner, E. W. (1974). Toward a more adequate conception of evaluation in the arts. Art Education, 27(7),
- Eisner, E. W. (1985). The art of educational evaluation. London & Philadelphia: The Falmer Press.
- Eisner, E. W. (1985). The educational imagination (2nd ed.). New York: Macmillan.
- Eisner, E. W. (1987). The role of discipline-based art education in America's schools. Los Angeles, CA: The Getty Center for Education in the Arts.
- Finlayson, S. J. (1988, April). Can we speak the language? Toward an understanding of standardized testing techniques with an examination of states' approaches to assessment of the visual art student. Paper presented at the annual meeting of the National Art Education Association.
- Fiske, E. B. (1987, December 27). U.S. testing of students raises growing debate. The New York Times. pp. 1, 22.
- Gaitskell, C. D. (1958). Children and their art. Methods for the elementary school. New York: Harcourt, Brace.
- Gaitskell, C. D., Hurwitz, A., & Day, M. (1982). Evaluating children's progress in art. Children and their art. (4th ed.) New York: Harcourt Brace Jovanovich.
- Gerhart, G. (1986). Effects of evaluative statements on artistic performance and motivation. Studies in Art Education, 27(2), 61-72.
- Greer, D., & Silverman, R. (1988). Making art important for every child. Educational Leadership, 45(4), 10-14.

- Greer, W. D. (1984). Discipline-based art education: Approaching art as a subject of study. Studies in Art Education, 25(4), 212-218.
- Greer, W. D. & Hoepfner, R. (1986). Achievement testing in the visual arts. Design for Arts in Education, 88(1), 43-47.
- Hamblen, K. A. (1985). The issue of technocratic rationality in discipline-based art education. Studies in Art Education, 27(4), 43-46.
- Hamblen, K. A. (1986). Testing in art: A view from the classroom. Design for Arts in Education, 88(1), 12-16.
- Hamblen, K. A. (1987). What general education can tell us about evaluation in art. Studies in Art Education, 28(4), 246-250.
- Hamblen, K. A. (1988). What does DBAE teach? Art Education, 41(2), 23-24, 33-35.
- Hausman, J. J. (1988). Back to the future: Reflections on present-day emphases in curriculum and evaluation. Art Education, 41(2), 37-41.
- Hodsoll, F. (1985). Some thoughts on arts education. Studies in Art Education, 26(4), 247-252.
- Hoepfner, R. (1984). Measuring student achievement in art. Studies in Art Education, 25(4), 251-258.
- Honetschlager, M. (1986). Statewide art assessment testing in Minnesota. NAEA News, 28(6), 3-4.
- Hull, C. (1984). Marking: A critical alternative. Journal of Curriculum Studies, 16(2), 155-164.

- Lai, M. K., & Shishido, J. (1987, April). A model for evaluating art education programs. Paper presented at the annual meeting of the American Educational Research Association.
- Lovano-Kerr, J. (1985). Implications of dbae for university education of teachers. Studies in Art Education, 26(4), 216-223.
- Lowenfeld, V. (1957). Creative and mental growth (3rd ed.). New York: Macmillan.
- Lowenfeld, V. & Brittain, W. L. (1987). Creative and mental growth (7th ed.). New York: Macmillan.
- Mehrens, W. A. & Lehmann, I. J. (1980). Standardized tests in education. New York: Holt, Rinehart, & Winston.
- Michigan Art Education Association and Michigan Department of Education. (1977-78). Analytical report. Michigan assessment tests in art. Statewide sample 1977-78. Lansing, Michigan: Author.
- Minnesota Department of Education. (1981). Minnesota educational assessment program. Secondary visual arts (Test). St. Paul, Minnesota: Author.
- Minnesota Department of Education. (1982). Minnesota educational assessment program. Elementary visual arts (Test). St. Paul, Minnesota: Author.
- Mitchell, J. V., Jr. (Ed.). (1983). Tests in print III. Lincoln, NE: University of Nebraska, Buross Institute of Mental Measurements.
- National Assessment of Educational Progress. (1977). Design and drawing skills: Selected results from the first national assessment of art (Art Report No. 06-A-03). Denver: Author.

National Art Education Association. (1986, December). NAEA News, p. 1-28.

National Assessment of Educational Progress. (1978). Knowledge about art: Selected results from the first national assessment of art ( Art Report No. 06-A-02). Denver: Author.

National Assessment of Educational Progress. (1980). The second assessment of art, 1978-79. Released exercise set (Art Report No. 10-A-25). Denver: Author.

National Assessment of Educational Progress. (1981a). Art and young Americans, 1974-1979: Results from the second national art assessment (Art Report No. 10-A-01). Denver: Author.

National Assessment of Educational Progress. (1981b). Procedural handbook: 1978-79 art assessment (Art Report No. 10-A-40). Denver; Author.

National Evaluation Systems, Inc. (1981a). Connecticut assessment of educational progress. Art and music. 1980-81 (Report no. BRPE-82-3). Amherst, Massachusetts: Author.

National Evaluation Systems, Inc. (1981b). 1980-81 Connecticut assessment of educational progress. Art and music. Technical supplement draft. Amherst, Massachusetts: Author.

Norris, E. & Goodwn, B. (1971). National assessment of educational progress: Art objectives (Bureau No. BR-8-0771). Ann Arbor, Michigan: National Assessment of Educational Progress.

Packard, S. (1984). Contemporary reform and the contents of curricula. Studies in Art Education, 25(4), 265-270.

Pennsylvania Department of Education. (1986a). Educational quality assessment. 1986 data. Harrisburg, Pennsylvania: Author.

- Pennsylvania Department of Education. (1986b). Educational quality assessment. Commentary. Harrisburg, Pennsylvania: Author.
- Pratt, T. B. (Ed.). (1983). National assessment of educational progress 1969-1983: A bibliography of documents in the eric database. Denver, CO: National Assessment of Educational Progress.
- Schubert, T. E. (1986). Art in education: Five rationales. Art Education, 39(1), 41-43.
- Shanker, A. (1988, April 24). Where we stand: Exams fail the test. The New York Times, p. E 7.
- Sims-Gunzenhauser, A., & Waanders, J. (1986, December). ETS and assessment in the arts. NAEA News, p. 8.
- Staff. (1986, December). States move toward testing in the arts. NAEA News, p. 1.
- Staff. (1986, December). Super sessions. NAEA News, p. 15.
- Staff. (1987, December). Indiana develops art achievement tests. NAEA News, 22(6), 3-4.
- Stiggins, R. J., Rubel, E., & Quellmalz, E. (1986). Measuring thinking skills in the classroom. Washington, D.C.: National Education Association of the United States.
- Tollifson, J. (1988). A balanced art curriculum makes sense. Educational Leadership, 45(4), 18-22.
- Utah State Board of Education. (1985). Art achievement inventory: Test A, B, C, D, E, F. Salt Lake City, Utah: Author.

- Utah State Board of Education. (1986). Art achievement inventory; Test A, B, C form I, Test A, B, C form II. Salt Lake City, Utah: Author.
- Ward, B. J. (1982). A look at students art achievements: Results from the national assessment of educational progress. Visual Arts Research, 16, 12-18.
- Wilson, B. (1971). Evaluation of learning in art education. In B. Bloom, T. Hastings, & G. Madaus, Handbook on formative and summative evaluation of student learning. New York: MacGraw-Hill.
- Wilson, B. (1975). National assessment and the future of art education. Peabody Journal of Education, 52(3), 213-216.
- Wolf, D. P. (1988). Opening up assessment. Educational Leadership, 45(4), 24-29.